FOOD AND DRUG ADMINISTRATION

CENTER FOR TOBACCO PRODUCTS

MODELING AND STATISTICAL METHODS FOR THE

REGULATORY ASSESSMENT OF TOBACCO PRODUCTS

A PUBLIC WORKSHOP

Friday, December 6, 2013

9200 Corporate Boulevard

Rockville, Maryland

This transcript has not been edited or corrected,

but appears as received from the commercial

transcribing service.

C O N T E N T S

C O N T E N T S (continued)

C O N T E N T S (continued)

C O N T E N T S (continued)

<u>P R O C E E D I N G S</u>

(8:30 a.m.)

Welcome and Opening Remarks

DR. DRESLER: So welcome back for our second day for our workshop on modeling. We're very glad to see those of you who stuck it out for the second day. I really liked yesterday, and I'm very much looking forward to today and some more good discussions.

Again, I think the hard questions -- one of them was directed to me yesterday, and I got out of it. Somebody else answered it for me. But these are really hard questions, so we do have to think about how to address. So I'm glad everybody is so engaged.

I know on your agenda it says 15 minutes for me to talk. I am not talking that long. So we are going to go ahead and get started with the first session. Dr. Ben Apelberg is the branch chief for our epidemiology branch, who you heard from yesterday, and he'll do this first session.

My short-term memory. There are

evaluation forms on the outside. So if you could

fill those out because that does help us in a

variety of ways. And if you just want to put them

in this box up here, that will help us. Thank you.

Moderator – Benjamin Apelberg

DR. APELBERG: Thanks, Carolyn.

Good morning, everyone, and I think it's

good that we start early, actually, because we have

quite a full day today. And hopefully everyone can

stay around to the end, to the panel discussion

about where we go from here.

So the title of this session is Models of

the Effects of Tobacco Policy, and we really have

some great speakers and panelists in this session.

We're going to start off with Dr. Doug Luke from

Washington University in St. Louis, and he's going

to give us, really, an overview of different system

science approaches, some of the different

methodologies that could be brought to bear on the

issue of tobacco and tobacco policy.

Then we're going to move to David Levy

from Georgetown, who is going to talk about his

SimSmoke work, which has been extensively published

and used to look at tobacco policy, both in the

U.S. as well as many countries around the world.

Then finally, we'll move to Dr. David

Mendez from University of Michigan, who's also done

extensive work in this area with his colleagues at

Michigan, looking at projecting tobacco use and

impacts of tobacco policies.

Then finally, we'll have a panel

discussion, where we're going to bring up two

additional panelists, Dr. Geoff Curtin from RIA

Services Company and Brian Morrison from Industrial

Economics, Incorporated. And I'll give everyone an

introduction.

So we're going to start off with Dr. Doug

Luke, and he's professor and director of the Center

for Public Health Systems Science at the George

Warren Brown School of Social Work at Washington

University in St. Louis.

He is a leading researcher in the areas

of health policy, organizational systems, and

tobacco control, and his work primarily focuses on

the evaluation, dissemination, and implementation of evidence-based public health policies.

So with that, I'll turn it over.

Presentation – Douglas Luke

DR. LUKE: Thanks a lot. Good morning, everybody. I was sort of hoping the music would continue for my talk. I think presentations are often more effective with a soundtrack.

I think we were asked to focus particularly on methods for modeling policy effects. And that's a lot of the work that I do. As Ben suggested, unlike some of the other talks that present maybe one model or one project, one approach, I'm going to be a little bit at 30,000 feet and give a number of examples.

The idea here, we heard yesterday, a number of people suggested the importance of diversity of methods. And so what I'm hopefully going to do is illustrate with a few examples the complimentary strengths of different modeling approaches, especially when we look at the effects of policy. And that's essentially what my overview

is.

At the end, I'm going to finish with what I think are -- here I'm calling them tensions and challenges. But they're really thoughts and observations, not just on the examples that I'm presenting, but some of the experiences I've had over the last few years.

In particular, as I've started reviewing more grant applications and peer-reviewed papers in this area, I'm starting to notice certain patterns that I think are interesting for us to think about. So that's where I'm going to leave us with.

Here at the beginning, I actually think I can speed through a few of these things for this audience. The methods I'm going to focus on are methods that are particularly appropriate for modeling of complex systems. And so why is it that tobacco control is a particularly appropriate thing for us to apply complex systems methods to?

I'm a social scientist by training, and so, when we're presented with theoretical models in particular, when social scientists say things are

complicated, they often organize it in ecological

models. This is a well-known example from Glass

and McAtee.

Here we see that health and health

behavior is nested within levels that sit above the

person or above the scan. And then we have things

going on underneath, and then there's a time axis.

Another example here, which comes from a

model of essentially social network influences on

human disease, and this is based on work of Lisa

Berkman and her group, we see that social networks

sit in between the broader social context and above

psychosocial mechanisms, behavioral mechanisms,

behavioral pathways, and so on. So again, we see

this tiered structure.

But more often lately when we look at

tobacco control -- well, I have this slide first.

So let's think specifically about how, when we're

talking about health -- let's talk about tobacco

control. And we've already seen a number of

examples yesterday.

But here we can see that there are

elements of tobacco control that exist at multiple levels, everything, as we say, from cells to society, from genes to global aspects. And Gary Giovino presented the host factor agent environment model, which also suggests this complex structure with different system domains.

This is what many of you are familiar with. This is essentially the causal map. This is the model that was presented in the ISIS monograph, NCI monograph. And this is a little hard to see, of course.

You can map the different domains of this system. So instead, here, with this picture, we don't clearly see a high, medium, and low level. We see different domains that are all interacting. And within each of these domains, you have very heterogeneous sets of actors. You've got consumers. You've got tobacco control scientists. You have state, national, local tobacco control programs. You have the industry. And all of these domains are interacting with each other in complicated ways.

So we want to study this. We want to model this. In particular, with policies, policies may come out of tobacco research. If they're implemented through tobacco control programs, we hope that they shape individual behavior.

Of course, communities and the industry respond to that. Government, in particular, for regulatory policy, they may be implemented in this section. This whole system has an economic component, and so on and so forth.

So again, what are appropriate methods for us to look at here? Well, as many of us know, traditional statistical methods have a number of weaknesses when dealing with complex systems. And when we look at the tobacco control policy system, we see a lot of these things.

So we see effects that are non-linear. Certainly the assumption of normality does not hold. I already mentioned that this is a system that has great heterogeneity, different types of objects, different types of people, different types of agencies, and quite clearly, a very dynamic

system.

This is from a social psychologist, Joe McGrath, and he called this a three-horned dilemma. And a lot of the traditional social science and public health science work ends up over here, in particular when we think of traditional clinical trials and laboratory experiments. And these types of studies are very good. Their emphasis is on the precision and control.

In modeling work, we actually end up over here, where in particular, the modeling is trying to get at some realistic aspect of a more complicated system so that the -- I think we had a question yesterday about which variables get controlled.

Well, control is not a concept that is quite as relevant when we're doing modeling because the point of modeling is to get all the important elements of a system in play. You don't want to remove the effects of important things. You actually want to model them and understand them.

So this all leads us to a different set

of methods that we might use. In public health, we

started calling these methods system science

methods. Different disciplines have different

terms for this. But three buckets of methods have

become fairly popular.

There's system dynamics and other

population-level modeling techniques. There's

network analysis; we saw a little network analysis

in Patrick Finley's presentation yesterday. And

there's agent-based modeling.

Late yesterday, someone talked about

top-down models and bottom-up models. And I think

that's a pretty good way to think about these

differences. System dynamic models, other

population-level models, work top-down, often.

Agent-based models work from the ground

up, where they start through modeling individual

agents and their behaviors, their interaction with

the environment. And sometimes these models meet

in the middle. We talked a little bit about that

yesterday when we were talking about hybrid models.

So the examples I'm going to go through roughly fit

into each of these three categories.

This table, which is from a paper we did I guess just a year ago, is an overly simplistic way of trying to suggest the different complimentary strengths of these methods. That is, the things that system dynamic models are good at are not necessarily the same things that agent-based models are good at.

This is really important. My sense, and some others of you -- Ross Hammond may know this better than I -- but in other disciplines, there's more of a competitive nature between modelers. I get the sense that the SD folk and the ABM folk duke it out a little bit.

Fortunately, I think public health, possibly because we're a little late to these methods, we seem to be a little more open-minded. And I saw that in the conversation yesterday, of being excited about the possibility of what multiple modeling approaches can do. So again, maybe that's a little Pollyanna-ish, but that's my sense.

So let's start top-down. Here we can talk about system dynamics models, but not necessarily just pure system dynamic models, also models that are essentially -- especially for tobacco policy, these are models that tend to take a population, a whole population, focus.

This is one example, a fairly recent example, from New Zealand. And this is a pure or full-fledged system dynamics model. What you see up here is just one part of it. This is a stocks and flows diagram from their model.

I'm not going to get into the details of any of these examples, really. But what I'm going to try to suggest are, again, what sorts of questions are these models well-suited for? Which helps us if we -- again, from yesterday's conversation, we talked about scoping and the importance of defining your starting question.

So if you're going to be starting with a question, you also, as a methodologist, need to think about what is the best method to match that question. So here, for example, with a system

dynamics model, these are forecasting plots that

look at -- I think this is cigarette consumption.

But it's four different scenarios over a

period of time, and it comes out of the system

dynamics model. So this is probably the prototypic

example of a population model used to forecast,

often either tobacco consumption or morbidity and

mortality, smoking prevalence, rates, and so on.

I'm not going to say too much. We've got

the two Davids coming right after me, and they have

great examples of population-level models. But

another thing, in addition to just pure

forecasting, is using population models to examine

counterfactuals.

So this is out of the SimSmoke projects

and Dave Levy's work, and I've circled here one of

the great things about these models is you can do

what-ifs. And you look ahead and you say, okay, if

nothing -- in this case, this is from Brazil. And

if nothing happened since 1989, what could we

forecast the smoking prevalence would be? And a

best-case scenario. I think we saw an example like

this yesterday.

But looking at the gap between these two numbers, it helps us make policy judgments about where we may end up, depending on what individual or sets of policies are implemented, in this case at the country level.

We also saw a little bit from CISNET. CISNET is a complicated set of models, but one of the things that comes out of these, as we said from Ted Holford's talk yesterday, again is a set of forecasting scenarios.

So here we see this is the forecasted. This is lung cancer deaths over time. And again, we see these counterfactuals of what would happen if we didn't have any tobacco control in the best case scenario. So again, a similar, very similar, idea.

I was glad to see where the conversation went yesterday. I included this example, which is from Tammy Tengs, and it's a different policy, tobacco policy, model. It's a forecasting modeling study, but what I thought was very useful in this

study is they present -- in this case it's more of

an econometric approach, and they're looking at

quality-adjusted life years when you invest in

tobacco control education.

This graph in particular shows that not

only are you forecasting, but you're looking at the

variability in your forecasting estimates. And a

lot of times, the policymakers are epidemiologic

colleagues. They want the point estimate that

comes out of these models.

Many times that's actually, I think, less

useful for us. Knowing what the range of possible

values are -- I think we talked about it in terms

of confidence yesterday. That helps us

tremendously. So, for example, if we're aiming for

a target -- let's say 2020 health goals -- and we

see with a range of possible predicted values that

we're not going to get near that target. Well, in

a sense, the best point estimate is less useful

than just knowing we're not going to be even in the

same ballpark.

Anyway, so the population models in

particular, they can do many different things.

What they've been used for in tobacco control is,

in particular, forecasting. One of the things I

think -- I may be misremembering. Was it Josh

Epstein that wrote the paper on Sixteen Things to

Do in Modeling Other than Forecasting?

So as I move to the next two buckets of

methods, think about that because I think in

particular, when we look at network analysis and

when we look at agent-based modeling, we'll start

seeing that there's different sorts of questions,

different sorts of goals in the modeling, that we

can use.

Network analysis is quite different.

It's, generally speaking, much more empirical than

SD modeling or agent-based modeling. But it's very

useful for studying and modeling the effects of

tobacco control policies.

This is one example from our work, where

we look at communication and collaboration networks

and state tobacco control programs. And this is

where we just use simple network mapping,

essentially, drawing the systems so that in this case what we were doing was identifying underlying organizational blueprints.

This is very useful for, in particular, CDC. They can use it to diagnose tobacco control systems; do they have the necessary components that a typical state tobacco control program would have?

Network mapping can be used not just to simply describe what's going on, but it can be used as an indicator or a marker of things happening in a policy space. So this is a fascinating paper from Denise Wipfli, and I think Tom Valente was an author on this as well.

This is in the international arena. These are two networks, and without even knowing a lot about network analysis, you can tell that something is different in this group compared to that group. This is communication. It's actually online communication between countries.

This is the set of countries that actually ended up adopting the international treaty of the Framework Convention on Tobacco Control.

These were countries that didn't adopt it. And
this is how they are talking to each other.

This is the use of network analysis as a
diagnostic tool to suggest there are really two
communication systems going on. In this case, it's
more in the political realm. But I think it's real
suggestive of the way that a network model may be
used.

One of the criticisms of network methods
is that they're purely descriptive. And the nice
thing, the exciting thing, about network science,
especially in the last 10 or 15 years or so, is
our ability to start building and testing true
stochastic models, that is, using network analysis
in a traditional modeling sense.

So just as one example of this, there's a
classic issue in networks, which is the issue of
homophily. We tend to be tied to others who are
similar to us. And the question that people have
been trying to figure out for a long time, and not
successfully, really, is how homophily arises.

So this is obviously a schematic, but it

shows, in the case of peer smoking or peer influence in smoking, what might be happening. So there's social selection, which is, here's a smoker who's looking around for friends and decides to become friends with other people who already are smokers. That's social selection.

Here is the other influence that might be happening, social influence. This is a nonsmoker who's already friends, but he or she happens to be friends with smokers. And over time, these friends influence a behavior change, so a person becomes a smoker. Both of these things are possible.

Until recently, we actually couldn't disentangle these processes until the development of these modeling technologies -- in particular, a set of techniques called exponential random graph models.

They're able to both -- it gets fairly technical, but they're able to control for things called dyad dependencies and allow us, with the appropriate study designs, to do what Elizabeth Mercken did here, which is, really for the first

time, show that -- in this case it's split by

gender -- but show that both these processes are

happening simultaneously. In a peer social network

system, selection and influence are happening.

Now, this is not a policy example, but

it's important for us when we start building

dynamic network processes into our models.

Traditionally -- well, before I get to that

example -- traditionally, we put network

information in a pretty simplistic way. How big is

the network? How many friends do you have?

Possibly a few other things.

But network science, we know a lot now

about network dynamic processes. And I think this

is a real important opportunity for modeling moving

forward.

This is just one last example of a

stochastic model where we're able to take what are

called multiplex networks. So this is a state

tobacco control system, and we measured contact,

just how often they meet and talk on the phone;

collaboration; and dissemination.

We were interested in trying to model the diffusion of evidence-based guidelines in a tobacco control system in a very important policy question. CDC wants everybody to know about best practices, but in fact, outside of the health department, out in the community, people don't know about these evidence-based guidelines.

What we're able to do, using these ERGM models, is essentially show that dissemination takes advantage of preexisting network ties. Again, this is a model. We're able to test it for the journal editors. We're able to get confidence intervals and p-values. And so they're very happy. So hopefully this suggests some network analysis examples that are useful for tobacco control policy questions.

I'm going to finish up with agent-based models. Now, there are fewer of these in tobacco control. We saw a great example from Patrick Finley yesterday in the Sandia group.

This is an example. This is from CISNET, and we heard about the smoking history generator

yesterday. I wanted to highlight this again. This is an example of a micro-simulation, I believe, where individuals are modeled.

So the modeling here produces a profile for individual people. And it gets aggregated and fed into the higher-level population CISNET model. I wanted to highlight this as a great example of these nesting of models that we were starting to talk about yesterday.

But agent-based models go a little further in that they model not just individuals but the possibility of individuals interacting with each other. So this is an early example in tobacco from Ross Hammond's group at Brookings where they were looking at different reactance patterns in a social network and then examining what would happen to a tobacco control message as it percolates through the network.

Depending on the social network structure, you see different patterns of communication. This is another example of bringing networks into a model, in this case an agent-based

model.

Let me finish up with work that we're just starting in our group. David Abrams asked about this yesterday. And Ross didn't have the graphs, but I do, so I'll present just a little bit about this.

The idea here is there are policies where we want to know what's happening, essentially, on the ground. We want to know how a policy is working. And people are starting to look at point of sale policies.

Here, in particular, we're wondering what would happen if you implemented policies that reduced the density of tobacco retailers. We already know there's an association between density and tobacco use. The question is, what would be the effects of a policy designed to reduce the density?

It's hard to do experiments with real communities on this. So we're creating a virtual town, Tobacco Town -- actually, a set of towns, urban, suburban, rich, poor -- and this is a

picture of our first essentially feasibility
prototype of the model.

The model looks a little bit like
SimCity, if you know that. But we're modeling
people who live in houses that go to school or
work. And on their way, once a day, if they need
cigarettes, they go and purchase them.

Then what we will be able to do is model
particular density reduction policies; for example,
remove retailers around schools, or remove a
particular type of retailer, for example
pharmacies.

What we're going to be looking for and
what we started looking for are, essentially, how
strong does the policy have to be. How much do you
have to reduce the retailer density before you
start seeing behavioral effects?

Like I said, we're very early on, but
we're starting to learn a lot of things already.
The most important thing, I think, is not only are
we identifying the important data that we're
bringing into our model, but we're identifying data

gaps. And again, David Abrams had a question about this.

For example, we know very precisely from surveillance systems who smokes, at what rate, when in their life course. We don't know how far people are willing to travel to buy cigarettes. That's a data gap. And things like EMA, ecological momentary assessment technologies, can help fill this.

So with our colleagues, they're starting to collect new data that we will then feed into this model. So there's a real iterative loop in model building with the tobacco control scientists.

This is an early graph that suggests there is a nonlinearity; that is, the density of retailers and the travel distance to purchase cigarettes, there may be a nonlinear relationship. And this is important for us because it suggests that a one-size-fits-all policy may not work, that certain types of communities start out here, certain types of communities start out here. For reducing density, the communities that start here,

we may need to go a lot further before we see effects.

I think I've gone long. So let me finish up with just some of the observations, which I've already hinted at a few of these. Let me actually start here.

I've already mentioned this. I think integrating social networks with modeling, either higher or lower, meaning population models or agent-based models, we're just at the beginning of that, and I think there's very exciting opportunities there.

I wanted to say something about this. I've been reviewing a number of things where it seems like no matter what the scientific purpose of a model is, people always feel that they have to put in the downstream outcomes in their models -- that is, smoking prevalence or mortality rates -- whereas in many cases what we're trying to figure out is not whether a policy works but how it works; that is, we're trying to use modeling to essentially reveal the causal mechanism. So, for

example, in Tobacco Town, what happens when
retailers disappear? What happens in the
community?

Now, if it's harder to buy cigarettes, we
will then eventually link it up to that reduces
tobacco consumption and that improves population
health. But not every model has to include those
downstream outcomes. And I think it's important
for funders. It's important for scientists to have
the courage of their convictions and not always
have a model have to include everything when it's
not always appropriate.

I've also started seeing people doing
things like look at end game strategies. So they
do a model that goes a hundred years in the future,
and they're reducing density -- in the example I'm
thinking of, they want to reduce the density to
95 percent.

But there's no dynamics. I mean, if
retailer density were reduced 95 percent, there
would be a response from retailers, from industry.
The whole system of tobacco marketing would have to

change.   But none of that is in the model.

So you have to be careful when you're
doing very ambitious models, that you always are
examining your assumptions, and at what point do
these assumptions change?

I think I'm going to stop there and
answer questions after -- Thank you.

(Applause.)

DR. APELBERG: Great.   Thanks, Doug.

Now we're got Dr. David Levy.   He's
currently a senior scientist for Pacific Institute
for Research and Evaluation and a professor of
oncology at Georgetown.   As I mentioned before,
he's currently overseeing the design and
development of the SimSmoke tobacco policy
simulation model, about which he's published over
60 peer-reviewed publications.   So we're glad to
have you here.

Presentation – David Levy

DR. LEVY: Thank you, Ben, and I'm very
glad to be here.   And good morning, everybody.

I'll now be talking about my baby, so to

speak, and that's SimSmoke. I've been working on it for well over 10 years now. In many ways it's stayed the same; in many ways it's changed quite a bit, and I'll talk a little bit about that.

Doug did a wonderful job of talking about the different types of models. SimSmoke is a compartmental model, and it's a unidirectional model. That is, policies affect smoking rates, affect death rates, and it moves in that direction.

What it does is it starts out with current, former, and never-smokers, and they evolve over time through initiation, cessation, and relapse. And based on that, there's a number of smoking-attributable deaths.

Now, the focus of this model is tobacco policies. That's always been the focus, the central part of the model, and these are the policies we've looked at: taxes, clean air laws, media campaigns, marketing bans, warning labels, cessation treatment, and youth access. These largely are viewed as demand reduction models. What we're getting into more now are supply-

oriented policies such as regulating the products themselves.

I want to directly move to policies. When I started out, I thought, okay. Well, modeling policies, this is going to be nice and easy. There's the literature. I'll just use that literature, and based on that literature, I'll make predictions.

When I started looking, I started out with taxes, and I felt pretty good. I moved on to youth access policies, and things got more difficult. And as I went through other policies, I found that there are lots of difficulties.

So what I tried to do, my goal was to not only look at the literature itself, but bring in theories, bring in economics, bring in sociology, bring in psychology and epidemiology.

Now, I'm trained as an economist. But what I did and what I think was most helpful to me is in having an expert panel. And on that expert panel were people with expertise in these different areas.

When I looked at a policy, my goal was, if you will -- and I'm putting this lightly -- is to tell a story that they could all buy into. That is, look at the evidence and try to get a consistent picture out of that. And I think that's one of the things modeling can do. It provides a structure for how we think about things.

Moving on to some of the specifics, when I model policies, if you look at the literature, what you find is most of the studies focus on the relationship. There's a change in policy and a Change in prevalence. And that becomes the basis of the effects. But how long- or short-term is that unclear? But nevertheless, that's where we get most of our information.

There are studies that look at initiation and cessation rates, and this is important. But usually the results are much less clear. And I think a lot of that results from the difficulties in measuring cessation and initiation, which is to say not that these aren't important, but we need to look very carefully at the measures and evaluate

those measures and use those measures in

conjunction with the other literature.

        The other thing which we tried to do is

look for age and gender differences. And something

that Luke mentioned is when you look at the

literature and you find very different kinds of

results is to try to then go back and think, why

are they different?

        What we're tried to do is develop

systematic models that explain the differences in

the results as much as the consistency. And

finally, to think about synergies. This is an area

where we only have very indirect evidence.

        Just one bit of background, and that's

distinguishing -- in SimSmoke, what we've tried to

do is model policies starting with a period before

policies are implemented and then allow for the

change in policies.

        We usually pick the initial period based

on when there's good data as well as trying to pick

a period before policies have been implemented. We

call that the tracking period, and we use that

tracking period not only to calibrate but also to validate the model and examine the role of past policies, such as I'll be showing in a minute. And then from a most recent year, we project forward.

Anyway, as I mentioned, I've been doing SimSmoke now for about 15 years, I guess, and these are the countries we've looked at. Now, I think, not only have we developed the models and we've tried to use the information we've gotten from the models, but I've tried very hard to document the results in publications.

I can't emphasize enough how important it is that we really get our work out there and expose it and get the reactions from other people.

I'm going to talk about three kinds of roles that I've had the pleasure to use SimSmoke for, somewhat overlapping. The first is advocacy, justifying policies. This was the original intent of the model. Second is planning, and third is heuristic. And what I mean by heuristic is not only understanding policies, but also finding gaps in the literature.

I'm going to start off with an example
from Brazil. With Brazil, first we validated the
model, and that's important because if it's not
validated, the confidence and results from it, of
course, are weakened.

The model did quite well. It did well
for the overall population. It did well for
subpopulations. We usually look at how well it
does by gender. We always have separate models for
males and females, and we look at, by gender, how
well it works for different age groups. By and
large, it did well for most of the age groups.
Usually where we do have the problems is at the
younger ages, which I think is not surprising.

Once we have a validated model, then we
could do the counterfactual. And essentially what
that involves is taking the model -- and the model,
of course, incorporates policies. And these are
what hopefully makes it predict well. But what we
then do is set the policies to the initial level.
For Brazil, the model started out in 1989, so we
kept policies at the 1989 level and then predicted

from there.

Here you can see the trend line under the counterfactual, where no policies change. And it predicts a slow decline in smoking rates, which is pretty much what you see in most countries that don't have very active policies.

Then we looked at the effects of individual policies. What I have here is for allowing price policies. And you see a big effect of price policies, which isn't unusual in many of the countries. And I would even go so far as to say in the vast majority of countries that have been very successful in reducing smoking rates, price increases through taxes have played a very major role.

Based on that, running the model for each of the policies, looking at the effects, we then come up with estimates of the percentage effects. And again, tax policies have the biggest effects. But other policies are very important here also, including smokefree air and media campaigns.

So once we've done this, we have the

rates under the counterfactual and the actual

rates. And based on that, we get a difference.

And based on that, I concluded that of the

reduction starting -- and this is overall smoking

prevalence -- in the decrease from 35.4 to

16.8 percent -- excuse me; this is for

males -- 46 percent of that decline was due to the

tobacco control policies, and those effects

increase over time, reaching as high as 60 percent.

Based on that then, I calculated the

smoking-attributable deaths under current policies

and under the counterfactuals, and that gave me for

each year the smoking-attributable deaths that

would occur.

Based on that, we accumulated over the

period, first from 1989 to 2010. And you see here

that we predicted about 5 million deaths under the

counterfactual versus 4.6 million with the actual

policies in place, yielding a difference of about

400,000 deaths. That's a lot.

But what you see is that the number of

deaths averted, also referred to as lives saved,

increases very rapidly over time. And that's how we got to the figure. By 2050, we've estimated that there are 7 million deaths averted as a result of the policies.

We've done this for many nations, and this gives you some of the ideas. And again, it's the in this case nations and states that had very active policies where you see the largest effects. And I think that's not surprising because when we look at countries, we look at the countries where -- and when people do studies, they do studies in countries that have been very active.

So, as a result, the model based on those predictions does best in those countries. Frankly, in countries where there is only one or two policies, the model doesn't predict as well. And again, I think that's not surprising because those are, if you will, the unpublished studies.

Now, one thing I'll mention here is one country in particular did better then predicted, and that's the United Kingdom, which leads you to question, well, why? Well, they have a very

extensive cessation treatment program. And that suggests that maybe that is an important component, something we can get from doing the modeling.

Moving on to planning, I omitted a slide here for the United States. But what we need to be aware of is different surveys give different results. When you look at the TUS and you compare it to the NHIS, there are different estimates that come out of it. And this is something that we need to be very aware of and think about.

I'm going to give you an example here from Ireland. In Ireland, I ran the model. There were two sources of data. I'd say, except for this one point here, it seemed to fit quite well. Okay? And so I thought, okay. Well, maybe that's a quirky point.

By the way, these are the largest surveys. There was one called the SLAN surveys, standard of living. And we had a point here. We had this point here. Then we had this point here. And I went and presented in Ireland and said, hmm, this doesn't look too good. I think there's some

problems with the model.

Then they said to me, well, I've got to tell you, we changed the way we did the survey, and we started including people with cell phones. And in particular, in Ireland, there was a large influx of people from Central Europe. So what we did then was we said, okay. Well, we could identify them. We can take them out. And we did that, and that gave us a point that was right on the line. It's the bottom green point.

So again, it's to point out the importance of looking very carefully at the data because that could have major implications on our results. So I can't emphasize enough the importance of using modeling to help us think about the data and the kind of information that we need.

Now, this is the more typical planning kinds of analysis. And again, I guess this is for Ireland. This is for males. And what we do is we run the model, and based on that -- and the important part here is the percentage reductions. We get ideas of the percentage reductions due to

policies.

Now, this is keeping in mind that Ireland had already implemented quite strong policies, including strong smokefree air policies. The only thing they didn't have is complete compliance, and that's why you see very small effects here from smokefree air policies. And taxes were also quite high. So it's not that taxes aren't effective, but they were very close to the 70 percent mark that NPower recommends.

But this gives us a ranking of policies, and also shows us how the effects vary over time. Some grow more than others. You see particularly large growth in the effects of tax policies, and that's because the biggest effects are on youth, and so it takes time for those results to work their way through the system.

What policies also make very clear is the effects of policies now are important, but the effects tend to be delayed.

Finally, when we look at policies, what I've found in my models is that, by and large,

using the best estimates of policies and even using

very optimistic estimates of the effects of

policies, you can't expect much more than a

50 percent reduction.

So what that suggests is that we're going

to need other policies, and that leads us to

explore other policies such as cessation treatment

policies, which I'll talk a little bit more about,

and also the potential importance of supply-

oriented policies, which affect the nature of the

tobacco product.

Now my favorite part of modeling, and

that's the heuristic part of it. I mentioned to

you earlier that one of the early policies I did

was youth access policies. And with youth access

policies, I had people with very different

views -- I won't go into their names, but with very

different views on the importance of youth access

policies and their potential effects.

I looked at the literature, and I found

that there were studies that suggested that they

had relatively large effects. This is on youth, in

terms of not only compliance but effects on smoking rates. And then there were other studies that found very small effects.

So we started thinking about how to develop this. And this led to a model where we brought into account -- now, youth access policies, for those of you not familiar, are geared towards reducing sales towards youth. But in talking to the people on my expert panel, I found that there were other very important sources such as youth gets cigarettes from parents, either stolen or given to them. They get them also from theft. And they get them from older friends. And so all these policies are potentially important.

So based on this, I developed a model to look at the effects of policies. And what I did was try to bring in the different aspects of policies and use these to come up with a theory of how they work. And what we developed was a multiplicative relationship; that is, you need to incorporate publicities, penalties, and compliance checks, and these will effect retail compliance.

What we also developed, again, based on the information I got from my expert panel, is an S-shaped curve, which is to say that you need to get compliance to very high levels.

Now, I did this model back about 10 years or so. But if you live in New York, you know that if there's somebody who's selling cigarettes, the kids are going to find out. It's going to be on the internet. And that's the common sense behind it.

Again, the modeling was a way to think about policies, and if I may go back, to think about not only the roles of the different policies, but also that publicity plays an important role. And this became important in my later modeling efforts, taking into account that the effects of policies are going to depend on attitudes of norms.

So I have built into the model synergies between particular policies, such as media campaigns, and other policies. And I think that's been an important part of explaining the effects of policies.

Another area where I think the modeling
has provided us some important intuitions is the
cessation treatment policy module. I built this on
a very -- this is a very standard decision tree on
the cessation treatment, where there is an attempt
to quit, and based on that, the individual can
choose different types of treatment.

Now, what's important here is this model
has also played an important role in thinking about
the synergies in policies because cessation
treatment policies -- and what I mean by that, let
me skip ahead, are policies related to
availability, towards financial access, that is,
subsidizing policies, quit lines, brief
interventions, web-based treatments, and follow-up
of care.

These are all policies that primarily, to
some extent, affect quit attempts. But more, they
affect the use of treatments. So based on that,
it's got me thinking about the synergies, and most
other policies affect quit attempts.

So what this suggests is that there's an

important synergy between cessation treatment policies and the other policies. And in my mind, that's what explains why Great Britain seems to be doing well, because they do have the strong cessation treatment policy orientation.

Recently, I have been focusing on other products, and in particular, smokeless, where the harm reduction becomes central. Briefly, the focus here was we need to look at the effects that the smokeless products will have in terms of moving people to these lower-risk products. But then we need to take into account whether or not they may lead to increased initiation and/or serve as a substitute for cessation, in which case they can be harm-increasing.

I just had a paper come out for Sweden that looks at these rates and models these over time. And what we found is that whereas the earlier studies found an important role of snus in terms of reducing smoking rates, in recent years it's much less clear.

So now we've begun modeling for the U.S.

And what I've found here -- and again, modeling forces you to think about these issues -- is in modeling for the U.S., we now have many different types of smokeless. We're going to need to take those into account. We're going to need to take into account this rapidly changing market and think about how these trends affect use.

This is potentially a very complex phenomenon and is going to require very careful thought about the effects of how these trends work themselves out.

The other important issue that we're going to have to grapple with is that policies, by and large, have been directed at cigarettes. And we have a pretty good idea of how those policies affect cigarettes. What we don't know very well -- there's been very limited study -- is how they affect the use of other products. We have some information from smokeless on how prices of smokeless versus prices of cigarettes affect the use of smokeless and cigarettes. But that's about it.

What we'll also need to do is look at policies not only that are directed at cigarettes, but that are directed more generally at all tobacco. And we'll need to look at those directed specifically at smokeless. And we're going to have to understand these better if we are going to really come up with models that have good predictive value.

The other thing, and this was alluded to yesterday, we have a pretty good idea of the risks of smokeless alone -- well, we have it for different products, at least. We have it for cigarettes. What we have much less information about is the risks associated with dual use. So this is another area that we'll have to give a lot of attention to.

Finally, while we're on heuristics, this is a diagram I have to indicate the kind of policy modeling that's been done. And as I've argued, most of the modeling has been on environment and/or on smoking behaviors. What we haven't done, and two important areas of the future, are going to be

to look at the specific role of the tobacco
industry and any feedback effects that has on
policies, as well as the effects on the
environment.

We're also going to have to look much
more at the individual genetics and that kind of
thing, and diet. I would add here that
socioeconomic status is an area we're going to have
to pay quite a bit of attention to in the future.

So this just is to give us an idea of
some of the links that we're going to need to
consider very carefully in the future. And with
that, I'd like to end with a plug for CISNET. I am
one of the CISNET members, so let me state that out
front.

But I'd like to point to CISNET as a
model for how we can proceed in the future. It's
going to be very important that we present our
results and compare the results of different
models. It's also going to be important that we do
things such as use common data sets, see how that
affects the results, and from that get ranges of

effects of the models over time.

So my concluding point is we're going to see many different kinds of models out there. They're going to serve many different types of purposes. The real focus, I think, in the future of modeling is to think about how these results either are consistent or not consistent with each other, and provide a unified framework for thinking about the different models and their results. Thank you.

(Applause.)

DR. APELBERG: Thank you, David.

Our final speaker today is going to be Dr. David Mendez. He's an associate professor at the Department of Health Management and Policy at the University of Michigan, School of Public Health. He's got extensive research experience in tobacco control as well as policies regarding residential radon. So we're pleased to have you.

Presentation – David Mendez

DR. MENDEZ: Good morning, and thank you for the invitation. I'm David Mendez from the

School of Public Health, University of Michigan,

and I'm going to talk today about a framework that

we've been using at the University of Michigan for

some time just to answer some questions.

The framework is going to be similar in

many ways to Dr. Levy's, and sometimes, especially

in policy details, less detailed than what he has

discovered, that he shared. And I want to

acknowledge that most of my work has been in

collaboration with Ken Warner from the University

of Michigan also.

So let's start with the framework that

I'm using in order to answer some, I think,

important questions. So I'm looking at the

population with smokers, former smokers, and never-

smokers. In this framework, the green circles are

never-smokers and the red circles are current

smokers, and then the yellow circles are former

smokers.

So those are individuals in the

population. We keep track of them before they

become adults, and then the model starts actually

keeping track of them -- we look at them before

they become adult, but the model starts keeping

track of them after they become 18 years old.

Then at that point, we just simply take a

lot at how many at 18 years old you have smokers,

and then we keep track about what's the

progression. And we don't anticipate that

there's -- or actually, the assumption is that

there's going to be very little initiation after

18, so we stop the initiation at 18, and then after

that, the cessation that we can isolate.

So if we are interested in the

interactions between those agents, between those

people, and if that's important to answer the

question, then we'll be looking at network

modelings and agent-based models. The questions

that we've started to talk about, think about, were

basically about prevalence. So what we did is just

consider that all the smokers are sometimes

homogeneous except for the transition from age to

older age, and the current smokers and never-

smokers and former smokers -- and former smokers,

depending on their years since quit, are going to have also some kind of similarity.

So then we just group those individuals with same characteristics into compartments, and instead of actually keeping track of them, we just count them. Right? We just figure out how many of them there are instead of actually looking at them individually.

Then we created this framework for analysis in which we started with never-smokers, and we get the transition up till they die. And there are some different policies that you can apply at the initiation and cessation rates of those individuals; and then after they complete their life cycles, we just take the difference between their survival curves and figure out the effects of them. So inside the model there is relative risk associated with former smokers and current smokers.

So how are the model dynamics? The model dynamics are -- we think the future is the present plus change, so smokers tomorrow are going to be

the smokers today plus some change. And the change is initiation minus cessation and death.

Individuals have some probability of initiation, and that's the inflow into the bathtub that we think the prevalence is. And then there's a probability of cessation of individuals that, depending on the age or different characteristics, that pushes people out. And that creates that bathtub.

So we are looking at the volume of the tank, which is the prevalence. But what we can influence is not the volume of the tank; we can influence the rate at which people get in and the rate at which people get out. And those are different things. Right?

So our output, the measure that we look at this prevalence. And prevalence takes some time in order to realize, given the changes in initiation and cessation. So if you do some simplifications of the model, a model like this, we can actually come up with a very unified solution for this model.

So I'm not now taking into account that
there are individual ages, but I just want to
figure out, at the very macro level, how this model
reacts, how the model performs.

If we assume that there is one single
initiation rate, and the cessation rate is stable
and mortality rate, what is important is that that
gives me the behavior of the model in time, and it
tells me that that prevalence is going to convert
to a steady state. And I can see what the steady
state is.

So if you tell me now the initiation
rate, cessation rate, or mortality rate, I tell you
what is that we can expect the prevalence to
stabilize to. And the prevalence is going to
actually try to catch that.

Also, what is important is that we know
that -- the parameters are sensitive to the output.
So actually, we can get variability of the output
because of variability of the parameters. So we
have data we can actually, with some very simple
assumptions, estimate or have an idea of what the

parameters of this model are.

So now I can disaggregate the model much more, and I can keep track of all the different ages and for former smokers, so to 30 years they quit, et cetera. And then I can use some data, and I did use some data, to estimate or calibrate -- but even more than calibrate, we created some confidence intervals for this.

We estimated cessation rates. And what we did is we used NHIS data from 1970 to 1993 by age group, and then we fitted that to the model projections. And we used the model as an estimated machine. And then we figured out that we estimated these six cessation rates before 1980, and after 1980 by different age groups.

Then we came up with this for 1970 to 1980, this negative from 18 to 30, because we are estimating net cessation. What it means is that there's some initiation after age 18 from 1970 to 1980, but it wasn't statistically significant, that parameter.

The post-1981, the post-1980, were

statistically significant and the fit was very

good, but also the likelihood. It was very

informative. So if we change those parameters, the

likelihood decreases dramatically. So there was a

lot of information in that matrix there.

It also gives us some -- it makes sense,

that estimation, that as younger people have lower

cessation rates, then there are more cessation

rates between 31 and 50, and there are more

cessation rate -- the cessation rate increases

after 50. So it actually makes sense. And we get

an overall 2.59 percent per year.

This is the fit of the model to the data

to that point, and this is the fit by different age

groups, which was quite good. For ages greater

than 65, the fit was a little bit more -- not as

good, but still quite informative.

So with that, we did some what-if

analysis. And the what-if analysis, we know what

the prevalence is going to come up with. But we

have now population changing. So on top of the

smokers coming in and out, we have anticipation of

population changing; that is, we put it into the model.

Then we just had some what-if analysis of what would happen if initiation rates will change because we saw that the model was quite sensitive to the initiation rates. Cessation rates, we tested cessation rates, the hypothesis that they have changed in the 1990s. And later I tested the hypothesis that it had changed in the early 2000s, and I couldn't find any major difference in cessation rate. But the initiation rates have been fluctuating a lot. So what happened is the initiation rates change; this is where we think the prevalence was going to be stabilized.

After that, then we kept track of the projections of the model with real data. And it has been quite good for some time, and very recently we are projecting about 19.1. For 2010, we were about 19.3. So the projections were quite good.

So we also expanded using CPS data. We used logistic regression with CPS-II data to fit

this relative risk for former smokers, and current

smokers and former smokers by age quit for male and

female, and that was incorporated into the model.

We also did some analysis to incorporate medical

costs into the model proportional to the relative

risk, and those were incorporated into the model.

These are some applications. One of the

applications that we had early on is taking a look

at the Healthy People 2010 targets and figure out

what -- in 1999, we figured out what are the

necessary changes in initiation and cessation that

needed to happen in order to achieve that target.

We found out that even if initiation

drops by 15 percent by 2010 -- which we are now at,

I think, 18 percent, so it did in 2013 -- cessation

rate would have to increase by fourfold in order to

achieve the -- at that point it wasn't 12 percent,

so I was working with a preliminary target of

12 percent.

So most interesting was that even if

initiation drops to zero, we still would have not

got the Healthy People target without increasing

threefold the cessation rate. Cessation rate

hasn't changed that much, so increasing threefold

was out of the range of possibility. So we gave

some idea of possible scenarios of what would be

possible and what targets we should be looking at.

Another thing that we did is say, well,

we don't think we're going to achieve Healthy

People 2010 target, but let's take a look at 2020.

What can we expect for Healthy People 2020? How

can we develop a better target for Healthy People

2020?

So we did the analysis of looking at what

is the region of the country which was performing

the best, if possible, or actually, that supplied a

lot of different tobacco control policies that are

showing in the prevalence.

So we figured out that Utah was the

smallest prevalence. But we used California, the

second lowest, because California has put in place

for many years some comprehensive tobacco control.

And then we used data from California to estimate

initiation and cessation rates for California, and

we did that with data.

Then we fitted that to the model to say,
well, if the country achieves -- and this is kind
of an exaggeration -- but if we continue to do as
we are going, what is status quo, this is what we
should expect. If the country can achieve the
cessation or initiation rate of California, then
this is what we should expect. So we're saying,
well, it's in the realm of possibilities. We don't
know what combinations of policies will make that,
but it happened in California.

So we did that analysis also of the
population in California to figure out that
they -- and actually, the analysis came up that it
wasn't just the composition of the population in
California, which was responsible to a lowest
prevalence. Two-thirds of the difference could be
attributable to tobacco control.

So we took a look at various optimistic
scenarios, and we figured out that by 2020, we can
be about 14.7 percent if we push the initiation and
cessation as California is. And then I think we

had a paper in which we were recommending something

close to 14 percent for Healthy People 2020 as an

achievable target. And then we did some analysis

just looking at where the country would be in

different scenarios.

So we recalibrated the model, and we just

did some what-if analysis of what the prevalence of

the country would be under different scenarios.

What is the status quo? What if we increased

initiation -- decreased 25 percent initiation and

increased cessation by 25 percent and 50 percent,

and so on, so to see what we think can be

achievable.

We used the model also, or a framework

like this, to figure out a managed care

organization, whether it's viable for them to offer

smoking cessation treatment. And we just build two

different worlds, simulation worlds, inside an MCO

and outside an MCO.

We put some parameters from a big MCO,

and we constructed an outside world, and we

interacted those two worlds. And we figured out

that with a lot of sensitivity analysis that is not
going to be financially beneficial for an MCO to
offer smoking cessation. So that was surprising
for us. Right?

So we figured out that that was going to
be beneficial for them. But because of the
turnover, what happens is they said anything that
they accrue is going to be lost, is going to be
gained by somebody else. So it's going to be a bit
externality effect. And so they don't have any
incentive to put that.

However, there is not a disincentive
also. So it kind of was a wash. So it's like very
good for society, but not a clear financial
incentive for an MCO. So that's a call for public
health there in policy, right, for a public health
measure. So we did a lot of sensitivity analysis
for that.

Also, we did some work figuring out what
is the impact of menthol on the population, menthol
cigarettes on the population. And the idea
is -- the experiment was, what if menthol doesn't

exist in the market? What would be the impact?

And with parameters supplied by the Tobacco

Products Scientific Advisory Committee from the

FDA, we created this scenario that look into the

status quo, what happens if menthol still exists in

the market, and the counterfactual, what if menthol

had never existed in the market.

Of course, there's never a clear way to

do that at the very beginning because we started

with some prevalence that includes menthol anyway

at the very beginning. So again, it's in the

conservative side.

So after a lot of sensitivity analysis,

we got a mean estimate of about 330,000 deaths

averted, or excess death because of menthol in 40

years, and 9 million extra smokers. Also, we did

some analysis of global smoking prevalence, so we

used the model to feed data from the WHO InfoBase

to about 80 countries that represent about 90

percent of the prevalence in the world, of the

global prevalence.

What we did with this model is we -- as

Dr. Levy alluded, it's very, very, very difficult
to find in the literature good data about effects
on initiation and cessation rates. But we used the
best what we could for initiation and cessation
rates of these NPower policies that can be applied
globally and affect initiation and cessation rates.

Well, the only thing that we didn't do
with initiation and cessation rates was the effect
of price that we applied to prevalence with price
elasticity.

So the question is, how do we combine
different policies? The model was that certain
policies affect the probability -- let's say
probability of quitting of a group of individuals.
Another policy affects the probability of quitting
of another group of individuals. And the question
is, when we have all these policies combined, how
will we just produce the combined effect?

So it's clearly not additive because then
we'll have more than 100 percent. So we went from
a very conservative scenario, and the worst case is
that the policies are totally correlated and means

that the best policy takes hold; they affect the same number of people, and the other people, they would have quit anyway, right, with -- just one policy is enough.

The next is the independence. So it takes a chunk, a proportional chunk, of the population. So actually, it shouldn't be a summation science, the multiplication science. So it's pretty much what Dr. Levy was alluded to, the multiplication of probabilities and find out a combined effect. So that's a best-case scenario.

Then we did a sensitivity analysis between the worst-case and best-case scenario, with a full Monte Carlo analysis to figure out what the impact is going to be with those policies. What we didn't do was look at the effect of one policy potentializing the others because we didn't have enough data.

So just looking at all this, fitting the model, we took several -- actually, a couple of years to do this analysis, fitting the model to the contrast for which we have data and before the

application of NPower policies, and then

extrapolating after that.

So we come up with an idea, the idea that

if nothing happens, we're going to see a very flat

prevalence in the world, about 23.7 percent to

22 percent because of different changes in

population. And that will happen because some of

the regions are going up and some of the regions

are going down in the world. But if we apply

NPower policies globally, so with a 100 percent

price increase, we have a rough estimate that from

23.7 percent, we are going to go down to 15

percent. We could go down to 15 percent in 2020

and 13 percent in 2010 [sic].

So that's an overview of the applications

of the model. One of the things that we are

working now is -- there are actually a couple of

things. The initiation and cessation rates are

pretty much exogenous to the model right now, and

they are affected by policy directly.

So what we are actually doing is testing

some mechanisms that are endogenizing the cessation

and initiation rates. The initiation rate depends

on peer pressure. So we have some feedback loop to

the initiation rate depending on the number of

people close to your age that smoke.

So we are testing different paths with a

system dynamic model that just produces a feedback.

But also, we are disaggregating a little bit more

with agent-based models, just testing a different

path of initiation pressure and cessation pressure

also.

So we can talk about that later. But I

think I've run out of time, so thank you.

(Applause.)

Discussion - Panelists

DR. APELBERG: Great. I'd like to invite

the speakers up, and our panelists for the panel

discussion for this session.

In addition to the three speakers, as I

mentioned, we have Dr. Geoff Curtin from RAI

Services Company. He's senior director of

regulatory oversight, he oversees the behavioral

research program for regulatory purposes; as well

as Brian Morrison, who's a principal at Industrial

Economics, Incorporated. He has been involved in

modeling and policy analysis for Health Canada.

So I'd just like to invite the new folks

here, the two panelists, to just offer any insights

or reactions to this initial set of presentations.

DR. CURTIN: Good morning. I'm Geoff

Curtin with RAI Services Company. I wanted to

first say I very much appreciated our being able to

participate in this conference. I think the talks

have been very good, and the discussion has been

very good as well.

I must say that this area of system

science is somewhat new to me. It stretches our

model, which I think you'll hear about in a few

minutes in the next session, our dynamic population

model, in some new ways. And it's thought-

provoking how we could bring other issues in beyond

the initial intent, which I think you'll hear about

as well.

I think from the talks that we just

listened to, it also points out or underscores that

a lot of these models were constructed for different purposes. So we had initially looked across the field of available models and wanted to ask a specific question, and I think we can talk about it, but it's a little different than maybe what we heard from Dr. Levy and Dr. Mendez.

I just bring it up because it has come up in terms of the model-sharing. I think it's great to have this type of diversity. At least with our model, it's really specific for FDA regulations and regulatory submissions. That's the way we look at it in that context. Quite a bit of resources went into it, and it is very specific in what it addresses.

But I thought the conversations and the presentations have been very informative, and I won't take up any more time.

MR. MORRISON: Thank you. I'm Brian Morrison with IAC. And I come to this -- well, it's actually the second visit that I have made to the Center for Tobacco Products. So in some ways, I too have been drinking from the fire hose the

last couple of days.

The first visit was to present a system dynamics model that we had developed for Health Canada to look at the implications and the potential public health benefits of a regulation requiring reduction in nicotine content in cigarettes sold in Canada to very low levels, which was largely an exploratory exercise.

To one of the points that Doug Luke made at the outset, we benefitted tremendously from the input of experts in the systems dynamic modeling arena, and particularly David Mendez and David Levy, who shared their insights with us at the outset of our modeling effort. So certainly I do see the collaboration occurring.

I also want to echo Geoff's comments about the importance of looking at issues from multiple perspectives and through multiple lenses. I think the value there is certainly in helping us to understand both what matters and, in particular, to the extent that we're evaluating public health benefits, what drives those benefits; but also, to

the extent that there is disagreement,

understanding the underlying causes of that

disagreement.

Ultimately, the modeling process, as part

of the policy development process, is an iterative

one, as we've heard over the course of the last

couple days. It directs further research and

data-gathering efforts. And certainly in the

context of the exercise that we were engaged in,

which was largely exploratory, it was very helpful

in identifying areas of uncertainty.

With respect to modeling for uncertainty,

in particular in looking prospectively at the

implications of new policy initiatives, I think

it's very important for the dialogue to extend not

just among the systems modelers but also to the

broader research community.

In particular, I believe Bill Poland

mentioned yesterday the potential use of subjective

probability assessment in the exploration of the

impacts of those areas of tobacco policy in which

there is great uncertainty about the implications

of a policy on initiation or on cessation.

In the work that we have done in the U.S. Environmental Protection Agency, certainly that is a technique that has come into wider and wider use. So there's much to be learned there as well.

But in general, this has been tremendously helpful to us. I greatly appreciate the opportunity to participate, and I'll stop there.

DR. APELBERG: Great. Thanks. Maybe I can just start up with a question, building off of, Brian, what you were talking about. And I think there's quite a bit of discussion, both today and yesterday, around this idea of uncertainty. And both of the Davids talked a little bit about in the realm of tobacco policy, the challenges of building off of the existing literature and applying that into these compartmental models.

Doug, you talked a little bit about hybrid approaches and the ranges of approaches to addressing those areas. I wondered if you had any insights on something that Dr. Finley talked about

yesterday, which is the potential for some of these

bottom-up approaches to be able to really inform

and fit into the macro-level modeling approaches

that you've been talking about, as opposed to just

building off of the existing published literature

around impacts of policies and changes in inputs

like initiation and cessation rates.

Anyone that wants to -- Doug, I'm looking

at you with respect to the challenges, maybe, and

opportunities for tying in different types of

models together.

DR. LUKE: I don't know if this gets

exactly at what you're aiming at, Ben. But one

thing I was thinking of as I was listening to all

of our presentations and this idea of what

different approaches can bring to the table,

there's a strong -- I guess in all three of our

presentations, the challenge of evaluating or

forecasting effects of policies when you've got

multiple policies in play.

Because again, traditional science would

view it as you test a single policy and you control

for everything else. We are trying to model

policies that -- even in a new policy, it gets

implemented in communities and places that already

have policies in play.

So to come back to part of your question,

why do we need different approaches? David, from

an SD perspective, you talked about sensitivity

analysis that has different assumptions about the

overlap in policies.

A very different approach through

agent-based modeling is you don't have to estimate

an interaction, essentially, in the model. You

don't have to go in with a parameter that estimates

the size of the additive or multiplicative effect.

Instead, you model the mechanisms of those policies

that are happening in the same, let's say,

neighborhood space, and you can see how the

policies operate together. Both approaches, I

think, will be necessary.

But I think that's one specific example

of it's really hard to figure out multiple

policies, and so we're going to have to model them

in different ways, and hopefully the results will
start converging.

DR. APELBERG: Thanks.

I wanted to open it up to the audience.
Any questions? Comments? Oh, sorry, David.
Please, go ahead.

DR. MENDEZ: No, no. I think I totally
agree. Particularly cessation rates, for example,
or cessation, we know now that cessation, because
of the Christakis model, that cessation happens in
groups. Right? So people don't quit isolated
Then there's the social network that cessation
probabilities transfer.

So it's very difficult just to come up
with a parameter in different situations just to
put it in the model. That's why, for example,
modeling specifically the social networks of
smokers and how they are grouped together now and
how they quit and how those different
policies -- like an opinion leader is quitting and
then the others are quitting -- so that's going to
be important in order to inform then -- it's a

micro-experiment that can be done with an agent-

based model, and then accounting for the macro

model and figure out what are the overall trends.

And more of that kind of interaction, I think, is

needed.

DR. LEVY: And if I may expand on that, I

think the point is that -- again, the main point is

that the microsimulation models can give us a

better understanding of the dynamics of how

policies work individually and together.

The added point I would make here is I

think that the microsimulation models can give us

some information, which we sorely lack, and that is

the expected range of results we might anticipate.

This is something we really don't have

very good information about. And I think the

microsimulation models, through the many runs,

could give us a better idea of that for individual

policies, and also potentially for when multiple

policies are implemented.

DR. APELBERG: Thanks. Geoff?

DR. CURTIN: Yes. Obviously, our models

don't focus on policy. But one of the things that

I recognized through these conversations is our

model is, if you will, complex where it needs to be

and simple where it can be.

So our sorting process is fairly complex,

but the user input and all that is fairly simple.

And we really wanted to ask and answer a simple

question, and that is, what are the potential

effects on the population with changes in tobacco

use behaviors?

The way the model was designed to keep it

simple is we weren't interested at the time in what

the motivations were that drove the behaviors, but

really the behaviors and what the outcomes were.

And I think the sense, from coming away from here

and what will drive our discussions forward, is,

some of the tools that we were thinking about that

would inform on potential behaviors or motivations,

could they actually be built into the front end of

the model?

For example, Patrick Finley's talk on

perceptions, or I'll call it outcome expectancies,

the literature would suggest that these can be
predictive of eventual tobacco use behaviors. And
from our perspective, trying to make estimations on
what might happen if, for example, a product is
given a risk modification order, that product is in
the market.

So being able to project is very
important for us in a different way than what we
heard today. So I guess that would be my takeaway,
is can we integrate some of the things we thought
would inform into the model more into the
model -- I guess you'd call it micromodeling -- and
how will that advance our understanding?

DR. APELBERG: Thanks. Any questions or
comments from the audience? We've got two over
here.

DR. POLAND: Thank you. Bill Poland of
Pharsight Corporation. I'm curious about the U.S.
models in particular because that's the most
relevant. And I'm curious -- in both the Levy and
Mendez models, are they fairly current or are they
being updated? And also, are they consistent?

So, for example, we saw a lot of nice detail on the cessation rates from Dr. Mendez. And I'm wondering if that's common across the two models or if there's different -- obviously there's different emphases in the models. But just curious about how consistent the data is.

DR. LEVY: I'll speak for my model first. In AAS, we do update it, and we try to keep it current, and we change it, and we try to make it increasingly relevant, to answer, I guess, your first question.

The second question is, David and I have been working, have had very different approaches, and we started modeling about the same time. And I've always been amazed by the similarity in the results that we've gotten. And again, these are totally independent efforts.

We both predicted slow declines in the absence of major policy changes, not too far apart from each other. We very independently developed menthol models, models of the effects of menthol, and our results are, in my view, remarkably close.

I think that that says -- it says that if we try to think about things and model them in a coherent manner, not only can we get ideas of the results, but we could get an idea of the range of results and we can, by looking at different models, get a better estimate of what might be affected. But it's very important to be doing different modeling and seeing whether or not the results are consistent.

DR. MENDEZ: So to answer your question, yes, I agree with David Levy's comment. We try to keep our model as current as possible. We actually keep looking at how the prevalence is doing and actually figuring out whether cessation rates have changed or not because of the model, given the way that we are estimating.

In the process, we can do better with the mortality data that we have because we use CPS-II right now and we can actually update that to better, more desegregated data. We are in the process to do that.

Also, I think we are getting the

prevalence of the U.S. down enough where these feedback effects are going to become a little bit more salient. So the idea of cessation -- again, we are looking at -- right now, cessation rate is stable. At what point are cessation rates going to show the dynamic behavior of a lot of people quitting? Right? So are they going to cascade?

So that's the kind of manipulation that -- there is a linear part of the model, it was insensitive for some time. But it's going to come to a point at which we are going to experience -- and hopefully experience some nonlinearities because there will be fewer and fewer smokers, and then there's going to be a much more cascading effect.

So that's the kind of things that we are updating the model just to -- the behavior is going to be the same, but it's going to try to incorporate the potential effects, I reckon.

DR. APELBERG: Did you want to follow up on that point?

DR. CURTIN: Yes. I just wanted to

follow up. And it's something David Levy said when he brought up the menthol modeling.

There was some discussion yesterday about how beneficial it would be if we used, if you will, the same nomenclature, if we identified smokers the same way, if we looked at abstinence the same way. And I couldn't help but think or recognize, listening to the talks the last several days -- and it actually goes towards this collaborative nature as well -- in that there needs to be some assurances going forward, or best practices, on how we actually use these models.

I come at it from a less sophisticated standpoint because I'm not a modeler. But the things I've heard over the last two days of running sensitivity analyses, determining what the most critical variable is in your model, looking back at that variable and saying, how robust is that variable? What was the data source? And are there ways to investigate that more thoroughly?

From our standpoint, we'd like to see that more universally done because it gives greater

confidence. And it would seem that being able to

compare across models that are diverse, while

following the same standards and being transparent,

that's the best way to encourage the scientific

dialogue on issues like menthol and other things.

DR. APELBERG: Thanks.

DR. SARKAR: I actually --

DR. APELBERG: Can you just introduce

yourself?

DR. SARKAR: Yes. My name is Mohamadi

Sarkar. I'm from Altria Client Services. I don't

have a question, more so a comment. And I'd like

the reaction of the presenters.

So when you look at the effect of policy

on the prevalence, for example, I would imagine

that it's important to consider the unintended

consequences from these policies, so, for example,

availability of counterfeit cigarettes and the

impact of that on initiation rates and access.

So without making the model too complex,

how does one consider some of the important

unintended consequences of changes in policy?

DR. LUKE: One of the things Ross Hammond said yesterday -- I think it was Ross who said, you look for surprises in your model, and they're either bugs or something important.

I think -- and I'm sort of new in this part of the modeling world -- but I think an unintended consequence is -- another way of thinking about that is something that comes out of the model that you might not have thought about before, and it makes you reexamine the model.

Again, something I think Ross said is you don't just say, ooh, that's an unintended consequence. You then examine the model, figure out what's going on, and see if that's something that informs you, something new about, let's say, the causal mechanism of the policy that you're examining.

DR. APELBERG: Brian, did you have --

MR. MORRISON: Yes. I think, too, there are times when you can anticipate unintended consequences, for example, in the context of the model that I was describing.

Clearly, in Canada there's a significant concern about the impact of black market activity on the effectiveness of any tobacco control policy. So that is something that can be explicitly built into a model and also tested through the same kinds of techniques that one would test other areas of uncertainty.

DR. MENDEZ: One of the things -- in the absence of data, you can model the mechanism of that unintended consequence and figure out, well, what is the level of that unintended consequence in order to offset your policy?

So you're saying, okay, my policy is producing this net effect. But then there's some unintended consequence -- let's say black market, et cetera. What would be the level the black market has to be in order to offset what I'm doing? Then you figure out, is that plausible or not? So that's one way that you can use your model.

DR. LEVY: I would add to that the emphasis that just the modeling -- we have virtually no idea what the black markets would be

if something like menthol was banned. We don't
have that. But we could begin modeling it. We
could look at experiences from as similar of
situations as we can find. And we could start
thinking about those reactions and incorporating
it, and at least come up with more systemic ways of
looking at it.

So when statements are made, well, this
is going to happen or that's -- when statements are
made, we can bring the conversation back to what
might be more reasonable expectations.

DR. APELBERG: Great. I think we're
going to have to leave it -- sorry. All right.
One more comment, then we're going to have to end
this and move on.

MS. HARTMAN: I just wanted to ask
you -- maybe this is naive --

DR. APELBERG: Oh, sorry. Can you please
introduce yourself?

MS. HARTMAN: I'm sorry. Right. Anne
Hartman, National Cancer Institute. This may be
naïve, but I assume more recently in the

literature, people have been doing more complex

analyses trying to look at combinations of various

policies. I'm not talking about the new

regulatory, but I'm talking about the ones that the

Davids and Doug talked about.

Can't you get some inkling about how

close the sensitivity analysis for interactions

might be from looking at the various policies and

combinations?

One of the things that come to mind is a

very simple analysis that was done a number of

years ago by Karen Messer, et al., where, over a

certain period of time when you could characterize

policies, they grouped states, like looking at New

York/New Jersey, that had a certain kind of

characteristic in types of policies they were

implementing, versus California, versus the

tobacco-growing states, and I imagine ITC, but of

course that's global as well as not just all in the

U.S.

But I'm saying have you used any of those

kinds of things to figure out what the joint

effects of policies are or how to partition it out?

Or have you really only been looking at single

policy analysis?

DR. MENDEZ: Well, definitely it's a

possibility, actually. That's a good point. What

I've done so far -- I've done three things so

far -- is look at places where they have

implemented a lot of policies and figure out what

the end result is. Right?

So I don't know how they interact, but I

know what the end result is, and I can say, this is

the end result. And I can disentangle, or try to

disentangle, what the interaction is, doing some

hypothesis or some modeling of what's going on.

But the overall effect I have.

Another way is doing the sensitivity

analysis that I did and actually go into some kind

of Monte Carlo analysis of risk analysis that go

from the worst-case scenario that I can think of to

the best-case scenario that I can think of and to

see whether that makes a difference. Right?

If by doing this range I cannot say

anything particularly important, then I need to get
a little bit more in that. But if by doing this my
results are still robust, then I have something
important to contribute.

The next thing -- and this is what we're
doing now -- is actually trying to model those
interactions with network analysis and agent-based
models to make hypotheses about how they interact,
how they work, and then figure out what the result
is, and try to put that into the model back.

DR. LUKE: I know we have to move on.
Very quickly on that, one of the things -- and this
touches on what Gary Giovino was saying I think
yesterday -- while we have very good surveillance
data in this country, knowing, for example, what
state programs are doing in terms of policy
implementation, it's almost impossible to figure
this out. There's just very poor data on policy
implementation, at least in the United States. And
modelers would love to have better data.

DR. APELBERG: Great. Thanks so much to
the panel. This was a really interesting

discussion. We're going to just roll into the next

session because we have a lot of really interesting

talks and discussion coming up, we hope.

(Applause.)

DR. APELBERG: But of course, there's a

water fountain right outside here as well as the

bathroom off to the right. So we'll continue to

go, but -- okay.

All right, scratch that. Let's take a

five-minute break since there seems to be a lot of

people migrating out. So we'll start again right

on time at 10:30. Thanks so much.

(Whereupon, a recess was taken.)

DR. APELBERG: This next session is on

Population Models Incorporating Multiple Tobacco

Products. And the moderator for that -- thank you

so much because it is easier not to yell -- is

Antonio Paredes, who is with our statistics branch.

Tony?

Moderator – Antonio Paredes

MR. PAREDES: Wonderful. So it looks

like we are pressed on time, so please, speakers,

let's try to keep to the original time that we

agreed upon, 20 to 25 minutes. And we'll go from

that.

So my name is Antonio Paredes, and I'm a

mathematical statistician and statistical reviewer

here with the Office of Science. And here, as you

see the title, we're going to talk about -- these

models that we're going to discuss in this section

are dealing with multiple products.

We're going to see two models

specifically dealing with multiple products, one

for Dr. Annette Bachand, and she's going to

introduce a model which is based on a Monte Carlo

approach via a Poisson model. And then we're going

to see Dr. Eric Vugrin. He's going to introduce a

model which is going to be based on a Markov chain,

Markov process approach.

Then Dr. Bill Poland is going to have a

discussion on topics that might be relevant to

developing this model. He's going to talk about

continuous versus discrete time systems and also

transitional rates versus Markov chains. And I'm

going to introduce each of the speakers here in a
second.

We're also going to have two panelists
that are going to join the discussion at the end of
the talks. And I'm going to ask each of those
panelists to introduce themselves and say a little
bit about their scientific background.

Those are Dr. Edward Boone, Virginia
Commonwealth University, and Raheema Muhammad-Kah
from Altria Client Services. So that is the plan,
so let's implement the plan now.

The first speaker is going to be
Dr. Annette Bachand, and she's an associate
professor, Colorado State University. And she's
also a senior science advisor for environment.

She has extensive experience working on
epidemiologic projects by providing epidemiological
and biostatistical expertise in analysis of
categorical and continual level exposure for
environmental and occupational sources in clinical
trials and retrospective studies.

Dr. Banchand has a PhD in public health

and biostatistics from the University of

Massachusetts, an MS in mathematics from the

University of Massachusetts, and a vordiplom in

mathematics and computer science from the

University of Tubingen in Germany. Dr. Banchand?

Presentation – Annette Banchand

DR. BANCHAND: Good morning, and thank

you. Today I'll be the one to switch the topic a

little bit, from one-product models to two-product

models.

As several speakers yesterday mentioned,

what's very important before we build a model is to

set a very specific goal. And the specific goal

for our approach was to build a population model

that incorporated exposure to two different

products and that was directly relevant to

regulatory assessment.

So we wanted it to directly address the

FDA requirements for MRTP applications. We also

wanted a model that was simple enough so that the

approach and the results could be very easily

communicated to non-statisticians.

The first question we had toward the end of this goal was, what features should our model have? Clearly, based on the goal, it had to model harm reduction. We wanted to incorporate the most commonly observed exposure histories. We wanted to make sure that any transitions would depend on age. We wanted mortality risk to depend on years smoked, years quit, and also on age. We wanted to make sure that we had a variability estimate; that was mentioned a few times also. And we wanted it to be validated.

The next question is, what was the approach that we were going to use? We decided to follow a hypothetical population over time. Again, yesterday one of the speakers mentioned the importance of looking at birth cohorts, and that's exactly what we did.

So our time variable is age categorized because of limitations of available data. And we don't have any restrictions with respect to the beginning age or the ending age or the age category width. So the user of our model can start and end

whenever he or she wants to.

We wanted to keep track of the exposure history of this birth cohort. So a transition being a change from one exposure state to another, and we wanted the user to put in whatever transition probabilities he or she thought reasonable.

In terms of mortality, we of course have a base case where there are only cigarettes available, and we calculate mortality rates in a nested Poisson model within our bigger modeler. But the input for the mortality rate is again user-specified.

So you can use a data set from any population you may be interested in. It needs to be stratified by age category, years of smoking, and years since quitting. And then you can feed it into our modeler, and the Poisson model will estimate the mortality rates.

For the counterfactual scenario for the cigarettes, cigarette part of it, of course the same Poisson model is being used. But for a

reduced harm product, we use excess relative risk,
which is basically a multiplier that reduces the
mortality rate for the MRTP.

Again, the ERR is user-specified, so you
might want to go with, say, an estimate from a
panel of experts. You might want to go with
extremes, something extremely high, extremely low,
for a sensitivity analysis.

How are mortalities calculated when you
have people doing different things throughout their
lifetime? So somebody may start as a never tobacco
user, then initiate smoking, and eventually switch
to an MRTP.

So for a person like that, our model
would use age-specific mortality for never-smokers
for the time when this person was a never-smoker.
It would then use duration-specific mortality for
current smokers during the appropriate time period.
And once the person has switched, we would apply
the mortality rate for former smokers -- because
the person is no longer smoking -- and the
mortality rate for current MRTP users.

Our model is implemented in WinBUGS and
uses Markov chain Monte Carlo techniques, which
give us variability of our output so we don't just
have a point estimate. As was mentioned earlier,
it is very important to have some sort of
confidence interval around the point estimate, and
this approach give us a posterior interval, which
takes care of that.

Now, we have quite a bit of possibilities
for our output. What you can get is the number of
survivors in the base case and the counterfactual,
and the difference between them. You can also get
the number person-years in the counterfactual and
base case, and the difference between them.

You can get the probability of surviving
from one particular age category to the last age
category, and the differences. And you can get the
estimated variability right now only for the number
of survivors and the difference between them.

Here is just one example, one use of our
model, one of very many possible approaches. In
this situation we're saying, okay, what happens if

we have a product assigned MRTP status and now some

people who would not have ever smoked cigarettes

actually initiate the MRTP?

So we put in smoking initiation and

cessation rates; in this example, they're for U.S.

males in the mid to late 2000s. And for this

example, we just, for simplicity, assumed that

among those who in the base case would never have

smoked, 5 percent of those now initiate the MRTP in

the counterfactual. We set an age range category.

We started with a population of a million; used

mortality input from the Kaiser Permanente study

and an ERR of .11.

What this chart shows you is that in this

situation, if we have 5 percent of base case

non-tobacco users initiating an MRTP instead, we

see about 3,000 fewer survivors in the

counterfactual compared to the base case, and that

is statistically significant. Zero would be no

difference between the base case and the

counterfactual.

But now what happens? Basically, what we

started out with was the unintended consequence of this MRTP assignation, that some who would not have smoked actually did start using it. But what about the intended consequence?

The intended consequence, or one of them, would be what if base case smoking initiators, people who would have actually smoked cigarettes, instead initiate the MRTP? Then what would happen?

So if about 13 percent of those smokers instead initiate the MRTP, the survival deficit that we saw initially is no longer statistically significant. If about 20 percent initiate the MRTP instead, we see no difference between the counterfactual and the base case. And if about 32 percent initiate the MRTP, we actually see a statistically significant survival benefit.

This is just one example of how our model can be used. And of course, you can turn it around, too, and you can say, let's start with the intended benefit. Let's see how many lives are saved. Now let's look at the unintended consequences, and let's see how bad they have to be

for the survival benefit to go away, and is that realistic.

What we have to keep in mind here, what's really important, is when we look at the never tobacco users and the proportion of those doing something else, we have a lot of people because if we look at smoking initiation in different age categories, it's usually around no more than 5 percent, probably less.

So the huge majority of people don't initiate smoking. And so even a small proportion of that is a very large absolute number; whereas if we then look at the people who would have initiated smoking, that's a much smaller group to start from. Therefore, the percentage must be much higher to get the same number of people.

In the previous session, speakers also mentioned that you have to do a ton of cessation to really make a difference compared to initiation, and that is just that there are so many more people available to initiate and there are a lot fewer people there that might possibly quit.

Also, over the past two days, we've heard a lot about validation of the model. At first we make sure that our sorting process work correctly, and we used Excel for that and found that, yes indeed, it did work correctly.

Next we looked at the predictive ability of the model. And of course the question is, well, can we use past data to predict the present? So for the base case, we decided to try to predict the 2006 U.S. life table for men using initiation and cessation data from 20 years in the past, to take latency into account; and for the counterfactual scenario, we used Swedish data because snus, which is an MRTP, has been used there consistently.

These are the results for the U.S. and the base case of the model. As you can see, these are the survivors in various age categories based on the U.S. life table, and this is what our model predicted. And you can see that we are extremely close.

For Sweden, again these are the numbers from the Swedish life table, and this is what our

model predicts. And again, we're very, very close.
So we were able to very accurately predict current
data using the base case and using the
counterfactual scenario.

Now that we have our model in place, the
question is, did we actually reach our goal? Can
we do with it what we have planned to do? And the
plan was to be able to directly address the FDA
requirements for MRTP applications.

Of course, that we can do with our just
basic original model because we can determine
whether it's reasonable to think that, okay, this
benefit will happen. Now there will be this
unintended consequence. What proportion of people
would have to do this for there to be a survival
deficit in the end? And we can use any combination
and address any questions that were asked.

But with small modifications -- again,
several people mentioned it's important that your
model, in a sense, has layers. You start with a
basic approach, and then you can go from there and
ask additional questions.

So one thing we can look at is, what
happens if the legal purchase age is increased? We
can also look at, well, what if the MRTP is
actually more attractive to youth? So maybe
overall initiation doesn't change, but maybe those
who initiate the MRTP instead of cigarettes maybe
start at a much younger age.

We can also look at removing a product
from the market. And of course, that's the example
of the menthol cigarettes. What happens if we have
a market with two products and then we remove one?

Now, what outcomes can we look at? Our
just basic original model looks at all-cause
mortality. We have a second version that looks at
cause-specific mortality. We're also looking at
morbidity, one simple approach just using
quality-of-life adjustments to our results. But we
have a second version that's under development that
actually looks directly at cause-specific mortality
within the model, and so the output is not
mortality but morbidity. We also are working on
another version that looks at the economic results

of different initiation switching, et cetera, approaches.

So of course, like any model, ours has limitations, too. We definitely had to employ several simplifying assumptions to make this model work. And currently, we only have posterior interval, so we only have a variability estimate for the number of survivors and the difference between survivors, not for person-years and not for the probability of surviving to the final age category.

A huge limitation, and that was mentioned a lot yesterday, too, is the availability of data. Clearly, the precision, the validity, of the input that is chosen by the user affects the validity and the certainty of the results. However, the model is extremely flexible. There isn't a single thing in the model that cannot be chosen by the user. So every single part of the input is chosen by the users. There are no restrictions on age, age categories, anything; you can model anything you want.

So we actually started calling this a
modeler rather than a model because it really just
is a tool, where the user can enter anything
they're interested in and see what they find.

What's also nice in the policy realm is
that the assumptions that were made that were put
into this modeler to get certain results are out in
the open. Every user who uses this model has to
very clearly state what went in to get the output.

Clearly, the model is really
comprehensive. We have accommodated what we
considered the most important exposure histories.
However, if somebody is just interested in
something very simple, they can use that, too, and
just put zero for the other transitions that
they're not interested in and just model that.

It's really nice and user-friendly
because if you want to do a sensitivity
analysis -- let's say you want to say, okay, I'm
using a certain set of mortality rates. I'm using
a certain ERR. And I'm setting all kinds of things
fixed. But what I really want to see is what

happens if I just change this one factor, or what happens if I just change these two factors?

You can do that. And you're not limited as to which factors you can test. So you can do sensitivity analyses on every factor and every combination of factors to see what happens.

Again, we have all the different applications that we can use the model for. And as was mentioned before, at this point we are not looking at why a particular initiation or cessation rate changed or a switching rate.

But this could definitely be a new step, something that would be added, basically, to the front of the model where we could say, okay, we have this policy change. Now how does this affect age-specific initiation switching, cessation, and then see what happens at the end. And so this is definitely something we are considering.

The validity, of course, is a strength. We have not only made sure that our sorting works and that our code is correct, but we have validated against actual population data.

Finally, what was most important to us, again, was the public health relevance. We wanted it to directly address the FDA requirements for MRTP applications, and that is exactly what it does. And so, again, we can take into account smokers who would have quit instead of switching to the MRTP. We can look at non-tobacco users initiating the MRTP instead.

We can look at a potential gateway effect. So people might be more likely to initiate the MRTP because it's a harm-reduced product, but then maybe use it as a gateway to switch to smoking. And we can test that against the beneficial transitions, which would be people switching to an MRTP instead of continuing to smoke, or people initiating the MRTP instead of initiating smoking.

So again, with our tipping point analyses, we can say, okay, if this happens, then how much of this other thing has to happen or these other two things have to happen for any harm or benefit to go away.

Before I hand over the microphone, one thing that was mentioned a lot recently was the sharing, making it simple to use, not just having a black box. And we've been working on a web tool where users can use the model very easily. And I think who these users will be is still in the open, but definitely FDA will be able to use it.

So other users can say, well, we didn't like what you used as your input. We think this other number would have been way better. And we can say, that's great. Put it in and see what you get.

So again, the goal of this was not to come up with specific results. The goal was to create a tool that can be used where people can have different opinions about what input data should be used, and put those in, and then see what happens at the end. Thank you.

(Applause.)

MR. PAREDES: Thank you very much.

Our next speaker is going to be Dr. Eric Vugrin. He's a member of the technical staff at

the Resilience and Regulatory Effects Department at
Sandia National Laboratories, and his primary area
of research interest is development of mathematical
tools and analysis method for complex systems such
as infrastructure, population, and nuclear waste
repository.

His work has been used by the U.S.
Department of Homeland Security, Department of
Energy, and the Environmental Protection Agency to
support both operational programs and policy
development efforts. He has a PhD in mathematics
from Virginia Polytechnic Institute and State
University. Dr. Vugrin?

Presentation – Eric Vugrin

DR. VUGRIN: Thanks, Tony. It's great to
be here today. Nice to be back in Virginia after
spending a little bit of time in the desert away
from Blacksburg and the East Coast. We got some
rain, so that's kind of foreign-looking to me.

But I'm going to continue in the vein of
what Dr. Banchand discussed and multiple products
and how do we think about modeling population

health when it's not just cigarettes that we're concerned about.

Whereas she took a cohort approach, I think the approach I'll take is a little bit more similar to the compartmental methods that Drs. Levy and Mendez used. But like we've been saying over and over again, the more models, more better. Right?

So how do we get here? Well, by this point in the workshop, there's been lots of talk about not just cigarettes but all sorts of different tobacco products. And over the last two decades, there's been a lot of change going on.

Some of it's policy, from the master settlement to the Family Smoking Prevention Act to the courts have seen banning of flavored cigarettes to specific packaging requirements. There have been changes in policy.

There are different products that are available. So cigarettes are still here. Snus is sometimes that's been introduced to the American markets. Electronic cigarettes are going on. And

not surprisingly, while all of these changes are
going on, there are changes in people's tobacco use
behaviors.

Cigarette use has slowly been decreasing.
Recently, cigar use amongst teens has been
increasing. And who knows what's going to happen
with electronic cigarettes. So with all of these
changes, it's reasonable to think that we may want
to consider a new set of population health models
to help compliment the existing approaches.

If we need new models, it's always a good
idea to think about what do we want out of these
types of models? First thing is the ability to
handle multiple products. We want to be able to
handle cigarettes. A lot of people have been
talking about electronic cigarettes the last two
days.

There's been a lot of movement, as
Dr. Banchand showed, to get two. Is two enough?
Do we need three? Four? I hope we don't need
dozens. But I'll more generally speak -- multiple
products is what we want. How far we go, I don't

know yet.

But when you get multiple products, that allows for new types of behaviors. You still have initiation and cessation that you think about, but there are behaviors like switching that you want to be able to represent, poly-use, and those things. And so as we increase the number of options, we increase the behaviors.

Additionally, I really liked Dr. Banchand's approach. We want this to be a flexible tool because we don't know everything that's going to be out there at this point. So we want the flexibility to consider different kinds of products.

Right now people are talking electronic cigarettes. Maybe that's something we need to do. There's been talk about smokeless. Something new might come out. So we don't want to cause ourselves to be too rigid, where we have to reinvent things over and over again. So let's potentially see if we can have the flexibility to incorporate different kinds of products.

We don't know what the risks are
associated with poly-use, necessarily. It was
brought up earlier. If we have smokeless relative
risks and we have cigarette relative risks, and
maybe we even throw in electronic cigarette
relative risks, how do they come together for a
poly-user?

So we want the flexibility to consider
different products that might have different types
of risk interactions: synergism, antagonism,
independence, those kind of things.

Cohort approaches are a good idea, multi-
cohort, single cohort. We also want to be able to
represent existing populations. We've heard
discussion about a national population health
standard as far as assessing the impact of policies
and products. So we want to be able to take an
existing population and see what's the impact to
that population if we go ahead and change things.

Additionally, dynamics. I think it's
very difficult to get a comprehensive look if
you've just got a static snapshot. Things take

time to evolve. Policies, products, can be shocks

to the system, and sometimes they manifest quickly,

but sometimes they don't.

So in sensitivity analyses that we've

done at Sandia Labs where we're looking at

initiation/ cessation factors and what's the impact

to, say, deaths and attributable deaths, initiation

doesn't start to have a significant impact until

further on in the scenario because it takes a

longer time for that mechanism to realize the

relative risks associated with the initiation

behaviors as opposed to the cessation behaviors.

Additionally, depending on what's your

analysis time period, you may end up with different

conclusions. So we want to look across a wide

range of times and see, okay, we might have one

result here, but we might have one result here, and

we might have another result here. Those are going

to affect the decisions that you make as far as

your timing goes.

So I mentioned it's going to be more of a

compartmental model that I'll does. We're going to

get a high level discussion, and we can get

nitty-gritty into the details outside of the

session afterwards.

But at the highest level, we start out

with a population of individuals that vary

according to state. I define a state to be a

unique combination of attributes, defined to be

sex, age, and tobacco product use. So if we were

thinking about a single tobacco product -- say

cigarettes -- the use options would be never,

current, or former.

Now when we get to multiple

products -- let's say two -- we can have never,

current, or former for each product. So now we're

looking at doublets. We have a never-never user, a

never-current user, a current-former user, and all

the possible combinations. If we talk about three

up to n products, then we're talking about n-

tuples. So it's a little bit more complex as far

as the tobacco use statuses that we're going to be

tracking.

At its highest level, it's a Markov model

of state transition and death. What that means is

that we're going to model transitions between

tobacco use state, statuses, and death as a

stochastic process. There's a probability

associated with changing your tobacco use and also

with dying, and so the probability of

transition -- this is where the Markov assumption

comes in -- depends only upon your current state.

The general idea is that we start out

with this initial population, and as we move

forward in time, we allow individuals to change

their tobacco use status. We allow them to die.

And so we're going to track this population over

time, and the deaths are going to affect the size

of the population. We'll also allow for births to

further supplement it, and ultimately migration.

What we're going to look for are similar

types of results to what were presented earlier

this morning. We want to know prevalences. We

want to know deaths. We want to know changes if we

implement policies, and those kinds of things.

So we've got some pretty basic

assumptions. We're only going to allow a state
transition once per time period. We typically look
at annual units, one-year periods, for our time
steps. And then if somebody has a current or
former tobacco product use status with regards to a
particular product, they may not go back to a never
status for that same product.

So what types of information do we need?
Again, it's very similar to the types of
information we saw earlier on. We're just adding
some additional dimensions as we move along. So we
need an initial population, which is the
distribution of states across the population on our
initial time. And we need transition
probabilities. So we need initiation, cessation,
relapse, switching.

But one of the differences here is
whereas in both cases when we're talking about a
cigarette-centric focus, and in this particular
case, a multi-product state, it's going to depend
upon the state. But in the multi-product
formulation, the probability that a cigarette

smoker is going to transition to, say, electronic cigarettes if we're considering three tobacco products -- say smokeless, cigarettes, and e-cigs -- the probability that a smoker transitions to e-cigarette use is going to depend upon not only their smoking status but also their smokeless status and also their electronic cigarette history.

So in this particular case, now we've got not just one individual parameter for initiation. We've got an initiation parameter for electronic cigarettes, for never, current, former, and all those different tobacco product use statuses again.

Also, our relative risks are going to vary by state. So if we're interested in all-cause mortality, as generally the types of things that we're doing right now, or we can focus on cause-specific, again we need relative risk. And relative risk is basically the probability of an individual with a specified state dying divided by the probability of an individual the same age and sex but never having used any of the tobacco products, is how I'll define the relative risk.

And the basic assumption is that use of tobacco
products inherently increases the relative risk
relative to never use.

All right. Director Zeller showed us how
complex things get as we think about more and more
products. And so when we look at tobacco use
transitions, just with a single product -- say
cigarettes -- we have three different states that
we typically think about -- never, current, former.
And I'm throwing all of the formers where we track
the years quit into one general bin here.

But we've got three different tobacco use
statuses that we think about, and there's three
different types of transitions that we think about:
initiation, cessation, and relapse.

So when we go to two products, everything
grows exponentially. Now instead of just never,
current, or former for the tobacco product use
statuses, we've got all the possible combinations.
So three times three gives us nine possible
statuses.

If we look at the possible set of

transitions here, where we're just looking at your

current state, dictating whether you can transition

or not, we've got 27 different possible transitions

here. We've gone from 3 and 3 now to 9 and 27.

I'm going to follow Director Zeller's lead and not

show you what three looks like. It just gets more

complicated.

So we have a few different options.

We've talked about the conceptual level;

numerically, how do we go about simulating these

and solving for the types of things that we want to

know? One is more of the typical standard Markov

formulation, where you do a microsimulation.

If we're looking at a population, you

start out with all of the different individuals

that you're thinking about. Numerically you flip a

coin to see whether or not they transition, whether

or not they die. And you track all of those

individuals through time as you move forward in

time. And what that allows you to do is explicitly

represent the stochastic events.

If you take more of a dynamical systems

approach, you can take those probabilities that

you're thinking about for transition and for death

and interpret those as mean rates of transition.

And now what we'll do, instead of tracking

individuals, we're going to track subgroups in this

population. And we're going to group those

subgroups according to a common state.

So if we want to know at some time TK

plus 1, how many people are associated with a

specific state coming from a different state, well,

what we do is we take the number in the previous

time step from the state that we're interested in

to see where they're transitioning from.

Then we say, what's the probability that

they're going to transition in that year, and

what's the probability that they don't die? We

multiply the population number times those

probabilities, and we get the mean number that

transitions from one state group to another.

Now, there's multiple states that you can

start out in and end up in the same one. So we're

going to have a summation over here, and so all of

the possible states that can transition into the
one that you care about. We're ultimately going to
add immigration numbers and birth numbers to give
us the total population in the next time step.

Typically, there are a lot of different
states that we're tracking. So if you think two
sexes, maybe a hundred different age groups that
you're worrying about, and if we're talking two
products, 27 different possible tobacco product use
status. If we go three, it's going to get even
more.

So instead of doing this individually,
we're going to create a matrix equation here where
we have an A matrix, which this one is our state
transition matrix. We've got a population vector
that we multiple it by to update how many people
come from the previous time step into this time
step. And now we've got immigration, and we've got
births coming along.

At this point in time, I'm going to talk
about how we use this. There was discussion about
models can be predictive so you can figure out

what's going to happen. Frequently, we're just
using this model to understand, what are the
possibilities? How can things change? What are
all the mechanisms and the different feedbacks that
are going on as we change one part of the system.
How do we see that changing?

So understanding is a big part of why
we're thinking about multi-product models right
now. And again, for simplicity, we'll stick with a
two-product model at this point in time.

Then we'll talk about an illustrative
analysis. This is notional. I'm not trying to
represent any possible real or fictitious -- real
or future products. This is hypothetical.

So we start out with an initial
population, and we take the demographics from the
2000 U.S. census in terms of age, sex, and
cigarette usage. And in 2000, we assumed that it's
cigarettes that we're only thinking about.

So we've got three years where we're
tracking this population where we're only thinking
about cigarettes. But in year 3, we have

introduction of a hypothetical new tobacco product.
For this particular analysis, we designate it to be
a lower risk one. And what that means is that its
excess relative risk is a quarter of what the
cigarette excess relative risks are for current
smokers, for former smokers, and all for the years
quit. That's going to be a quarter of what the
values are for cigarettes.

So with the introduction of this new
tobacco product, there are a couple of different
behaviors and things that are relevant to this
introduction. First off, now we can have switching
from cigarettes to the new product.

Current smokers, they may choose to
continue smoking. They may choose to just quit.
Or they may choose to quit and pick up the new
product. But they've got a new option here with
switching. And also, current smokers could pick up
the new product and use both cigarettes and this
new product. They've got poly-use in this case,
dual use. So these two impacts are really going to
affect the current smokers relative to a baseline

scenario, where we've got no alternative product.

The second set of behaviors that we look at are initiation of the new product amongst the never-smokers. So these are people who are not currently using tobacco products, but now they have the option to.

So some of the never-smokers will pick up this new product, and they may stick with it. They may quit. They may quit it and switch to cigarettes. So this is now also an alternative pathway in which an individual could end up ultimately smoking cigarettes, or they could just stick with the new product.

I really liked what Dr. Hammond said yesterday about results are one thing, but understanding why you get those results -- it's really important to understand.

So what we do is even though the scenario we're looking at has all of these next complexities relative to a baseline scenario where there's only the single product, what we're going to do is we're going to impact these new types of behaviors and

stages so we can see what the individual impacts of
each of these different kinds of behaviors are.

So in blue, we'll have the baseline run,
where we've got no new tobacco product. We've just
got cigarettes in the system. Then we're going to
have the impacts to current smokers, where we allow
switching from smoking to the new product. And we
also allow smokers to become poly-users.

But we're not going to have new product
initiation among never-smokers. So this is going
to see, as current smokers switch, what's the
impact relative to the baseline?

Then we're going to have one where we're
really going to just focus on the never-smokers.
So we're not going to have current smokers
switching, but we're going to have never-smokers
potentially initiating this new product and then
ultimately possibly picking up smoking in the end,
or just sticking with what they've got, or
quitting. But we're not going to have current
smokers switching or doing polys.

Then ultimately, we want to see the

combined effect of all of these different

behaviors. So in the cyan diamonds, we've got our

overall impact.

So the first metric that we look at is

adult smoking prevalence. Up here we see this is

when we introduce the new product. And so this

blue line right here is our baseline value. And we

see that if we only introduce switching and polys

amongst cigarette users, we get a drop in cigarette

prevalence. Not surprising. You've got some

people who were previously smoking. Now they're

switching.

When we allow the initiation is really

not a significant impact relative to the baseline.

We do get a slight dip initially because those new

initiates for the new product have to come from a

group, and that some of those people that pick up

the new product would have otherwise started

smoking cigarettes. Some of them would have never

started smoking cigarettes.

So initially, we get a little bit of a

dropoff. It's really small, so squint your eyes

and you might be able to see it easier. But by the

end, we see that some of these new product users

that would have started cigarettes but

didn't -- they chose the new product -- some of

them end up smoking cigarettes. And so we really

don't see a significant change from the baseline.

So under this particular set of values

for this scenario, and it's notional, we see that

ultimately the impact in this particular case,

adult smoking prevalence drops relative to the

baseline value.

So if we want to see what's the impact in

terms of changes and death, what we see here is

that this is the relative risk -- excuse me. This

is the impact of switching for the current smokers.

And we see that, not surprisingly, when they switch

from cigarettes to a lower risk product, that we

get a decrease in the number of deaths.

However, if we just have the new product

initiation going on, we see that ultimately we do

get an increase in the number of deaths because

even though some people are picking up the new

product, which is lower risk, we have on the next
slide, ultimately, an overall increase in the use
of tobacco products.

So tobacco control is not just about
cigarettes. So we'd want to see what's the overall
tobacco use. And in this particular case, we've
got the baseline down here. The overall impact is
that we have an increase in the overall tobacco
product usage.

Even when we allow the switching to go
on, the reason we get an increase in the overall
usage is that some of the people who quit
cigarettes and take up the new product would have
quit otherwise. And so, at least in this
particular formulation, we get an increase there.

So we've used point estimates for all of
these runs to show things relatively simply. But
we've heard many times now, uncertainty matters.
And there's a whole lot of different ways to look
at it, especially because we don't necessarily know
everything about the future. My kids do, but I
don't.

So what we're going to do is we're going to show just some very quit results. What if we include uncertainty? And so we can look at what if we change an individual variable? But as Dr. Finley mentioned yesterday, multiple variables and considering the uncertainty associated with those and the overall impact is important, too.

This is one of the results. This is the overall impact scenario that we're looking at. And what we're going to do is we're going to vary the value of the new product initiation rate. And so here is when we have the lowest new product initiation rate. And so we've still got smokers who are quitting, but amongst the never-users, right now that rate is set to zero.

What each of these lines show is these are the contour of distributions of decreases in death that are impacted when you put the distribution of possible new initiation rates -- initiation rates for the new product.

So we see, for a lower initiation rate, that leads to more decreases and death. And at

30 years, for all of the variables we're

considering -- this is no change relative to the

baseline -- all of the values we considered

resulted in a decrease.

However, remember, initiation takes a

longer time to work through the system. So if we

go even farther out, now we're able to see the

impacts of increased initiation. So now here's it

not all values are leading to a decrease in the

number of deaths, but there's a threshold which, if

you're above this value, you get an increase in

deaths; if you're below, you get a decrease in

deaths.

In this particular case, when we do an

individual single-parameter sweep, we see, okay,

the conclusions are not so clear now. And in this

particular case, if we do a multivariate Monte

Carlo simulation with a Latin hypercube design,

what we're going to have is a whole bunch of

different possible parameters here.

We're going to potentially vary excess

relative risk, the rate of switching product, how

many of the switchers are coming from would-be
quitters and how many are coming from those who
would have continued smoking, nine different
factors here.

When we plot the means plus standard
deviation 1 and 2, we see in this particular case,
uh-oh. Okay. Now we get a different result.
We've got an increase in deaths associated with
here as far as the mean goes, and then we also have
to see, if these are the standard deviation bars,
where do we end up here?

Do we feel more confident? I don't know.
But we have more information. And so what the
multi-product model helps us do is it helps us
think through what's going on. So it's a complex
problem.

I know I raced through the results. But
what we were trying to show is the variety and the
richness of results and conclusions, one might
potentially come up as you think about multiple
products. It's not an easy problem. Probably
formulation is the easiest part. Interpretation

and analysis are going to be difficult.

I'm just about out of time. But I'm going to say in my opinion, we need to start thinking about multi-products and how they can be used to assess policy in a changing marketplace. We've got changing behaviors, policies, marketplace. Multi-product models can help us think through this.

I think that there's been great work done in the work of Drs. Levy and Mendez, which are more cigarette-centric. I'm glad to see that they're expanding into menthol and the smokeless ones. And their work provides a great springboard for thinking about multiple products simultaneously.

I think the approach here is relatively flexible. We like to use it. We have a lot of different ways we can pull and move upon it. But there's still a lot of hard work to be done. We didn't talk about how do you come up with the cessation values or the initiation, all of the probabilities that you want.

So it's going to be difficult. There are

going to be data changes. We've heard it's hard

for just cigarettes. I don't know if it's harder;

I'll say it's definitely more work for multiple

products.

        But as we think about this, we need to

also get a multi-product mindset. If I know

everything I need to know about cigarettes and I

know everything I need to know about smokeless, do

I know everything I need to know about the combined

system? It doesn't always translate that way.

        With that, I'll wrap up. Thank you.

        (Applause.)

        MR. PAREDES: All right. Now we're going

to have a general discussion in some senses on

system dynamic. Our next speaker is going to be

Dr. Bill Poland, and he's a vice president and lead

scientist of Pharsight Consulting Services.

        Dr. Poland has provided industry guidance

and strategy, decisions through scientific and

decisional-related modeling, and simulation since

1994. Over the past 15 years, he has been at

Pharsight, and he has provided modeling and

recommendations to the pharmaceutical industry in over 60 projects, including regulatory submission material and modeling in response to regulation query.

Dr. Poland has a PhD in engineering economic systems from Stanford University and also a master of science from Stanford University, and master of science in operational research from USC Berkeley, and a B.S. in engineering from Harvard University. Dr. Poland?

Presentation – Bill Poland

DR. POLAND: Thank you, Tony, and thanks, everybody. I know we have two more talks before lunch after mine, so I'll see if I can get through this efficiently.

I'd like to switch gears a bit and talk not about some nifty model; we don't have one to show you yet. So I thought I'd step back and give you a little bit of the theory that lies behind these models; in particular, look at the questions of does it matter if we discretize time and take time steps or leave time continuous. And does it

matter if we actually model the transition steps as probabilities versus deterministic transition rates?

So I'd like to tell you what I'm going to talk about, and then say it, and then tell you what I said. Maybe the last of those, I'll do it very quickly due to time constraints.

I'm going to start just by asserting that continuous time is the natural view. Hopefully I don't have objections there. But models can take discrete time steps with no loss of accuracy.

Discrete time modeling has the other advantage that it's simpler to simulate. You can just step through time. But either way, whether time is continuous or discrete, we can express the solution in terms of terms that are exponential in time and that approach an equilibrium. Now, if conditions are changing, that equilibrium will be a moving target. But at least it's a target that we can calculate.

Then I'll just briefly talk about Markov chain models, which use individual transition

probabilities per period rather than rates or proportions of people that make a transition in each period, and argue that they are fully consistent with deterministic models.

Now, the most important question that we face is, is the right term "dynamic systems" or system dynamics? And I'd argue, actually, that the difference is just that system dynamics is the application of dynamic systems to complex problems or systems.

Now, I learned dynamic systems from an old text that I'd highly recommend, Luenberger, 1979, A Direction to Dynamic Systems. And one reviewer called it "the best mathematical textbook ever written." So I'd recommend that one.

He talks about continuous time and discrete time. And the idea of continuous time is that the rate of change of your vector of variables x is a function of your current state and other time variables. In discrete time, we just look at time t plus 1 and we say that state there is a function of the system at time t and other time

variables.

One thing to note is that the units are different. So be careful when you're looking at rates. In a continuous time system, those are probably per year, whereas in a discrete time system, they're going to be in units of people with the time already built into the periods.

Now, discrete time modeling is convenient when events or consequences either occur or are accounted for only at discrete time points or periods. So we can just step through time. And with continuous time systems, we might have to do numerical integration with careful attention to the size of the steps. So we avoid that with discrete time.

Let's take a look at quick examples. I love Dr. Mendez's tank and pipe figure, which you already saw this morning. So here, smoking cessation is going to be a rate which is proportional, perhaps, to the level of the tank. So we might have the rate of change of the prevalence axis minus kx, so the bigger x for the

tank level is, the bigger the flow rate.

An example of a discrete time system, which I think is important, is your basic cohort population model, which I think is worth taking a quick look at since it's so important in all the work we've seen.

We have age cohorts that don't have to be a single year, although I think the models we've talked about do use single year steps. And we have time periods, which will be in concert with the age cohort. So if we have five-year cohorts, we'd have five-year time periods.

We have survival proportions. We have birth rates -- we call them beta i and alpha i -- and our state is x sub i, by cohort of t. And so we just have two equations, a survival equation; state time t plus 1 for cohort i plus 1 is just the state of cohort i at time t times their survival proportion. Notice I didn't call it a survival rate because I'm trying to emphasize that it is not a per-unit time measure.

Then we have the birth equation, which is

cohort zero; x0 of t plus 1 is a function of the

levels of all the cohorts times their birth rates.

So they'll be zero for the first and last cohorts,

but positive numbers in the middle.

Good. So now I want to hit you with a

little math because there's a couple of really neat

results that are behind the more complex models

that we've seen. And I'll start with just a single

equation because the math is a little simpler, and

then I'll show this where x is a vector.

So there's a wonderful analogy here. In

continuous time, dx/dt, we'll say, is equal to a

proportional transition rate times x plus a

constant growth rate, in the simplest model. Of

course, we can make that a function of time in more

complex models. The discrete analogue is just

about the same except instead of a rate, we have a

survival proportion. We have an initial condition.

And I've noted, again, that the units are

different.

The first thing that I like to do

whenever I see a differential equation model or a

difference equation model is ask, what's the

equilibrium? Because it's so simple to calculate.

You just set dx/dt to zero, or xt plus 1 equal to

xt, and solve. And you can probably do that in

your head and make sure I'm right here. It should

get b over k, or b over 1 minus p on the right.

It turns out that the solution in general

for x over time can be expressed as a weighted

average of the initial condition and the

equilibrium condition, where the weight is actually

an exponential function of time, 1 minus e to the

minus kt, which approaches 1, or 1 minus p to the t

if we're in discrete time, which also approaches 1.

So that's neat result number one.

Neat result number two is that we can

make these completely equivalent if we just set the

equilibrium values the same and the b's the same,

or set the equilibrium values the same, I should

say. And we get k equals minus the log of p, or p

equals e to the minus k, and that expression for b.

So let's just see what that looks like in

a little example, where k here is minus 1/2 and b

is 50. So the equilibrium is 50 over 1/2, which is

100. If we start at equilibrium, we'll stay at

equilibrium. But if we start above equilibrium,

we'll asymptotically approach equilibrium.

Similarly, if we start below, we'll asymptotically

approach it.

Now let's do the same model in discrete

time with the changes to the rate and the growth

constant. And it's looking pretty similar. We

only have predictions, of course, at discrete time

points. And indeed, it's a perfect match, so

there's no loss of accuracy when we went from

continuous to discrete time. And that's a

reassuring result, I think.

Now I'd like to just generalize to many

variables. But don't worry, this slide is just

about the same as before. I've changed notations

slightly so that what was minus k is now a matrix

A. We have a matrix P; x and b are vectors. And

these expressions are completely analogous, and so

is the solution.

So that's great news. And I'll just

mention as an aside, for those who have a

pharmacometrics background, this should start to

look familiar because in pharmacokinetics, we model

concentrations of a drug coming into the body and

getting distributed to different compartments.

The drug goes in and out of peripheral

tissues, sort of like going from current smokers to

former smokers, and relapsing back. And we know

that the drug concentrations take a while to settle

to equilibrium. They change at different rates

before they approach equilibrium.

So one consequence of this generalized

solution, which is now a sum of exponential terms,

is that you might not approach equilibrium right

away. It might bounce around a bit first.

So to illustrate that, let's take a

little example, which is just set up in a

spreadsheet where we have -- and this is just

illustrative -- have a million youths coming in,

and they're spread between never-smokers and

smokers at an 80 to 20 ratio.

They're becoming former smokers at 1 and

a half percent per year with no relapse, just for simplicity. And then everybody dies in the long run, the smokers at 1 and a half percent, and the former smokers at 1 percent, and the never-smokers lower at .7 percent.

So those numbers are captured in the A and B matrix, the A matrix and the B vector. We also need initial conditions, so we have x0. In this example, I set the never-smokers twice as great as the smokers; maybe there was a recent change in policy that reduced the initiation rate so that we're off of equilibrium.

But we can now immediately calculate the long run equilibrium, all else staying the same, as minus a to the minus 1b. And it comes out that the smoker proportion will be about 5 percent. So that's actually surprising already because we saw an 80/20 ratio up here. But it's because of the cessation that's bringing down the long run proportion of smokers.

How long, though, does it take to get there? Well, we can simulate out, and here's a

plot showing current smokers in red, former smokers in the grey area, and then the never-smokers are the rest of the population plotted on the right axis. And it takes centuries to reach equilibrium.

The other interesting thing is that the number and proportion of former smokers actually increases for a century before it starts to decline. So the dynamics can be very slow, and that's important to remember, especially if you're affecting initiation rates versus cessation rates, which are more immediate.

But in order to make the model useful, we need to expand it to incorporate lots of other variables. We can put in dual product users. We can put in transitions between former users and dual product users, test product which might be, say, e-cigarettes or this might be cigarettes.

We can put in all the transitions; in this case, I didn't put in a transition from former users straight to dual product users. You could if you wanted. But I recommend simplifying wherever you can.

You can put in explicit experimenters. And most importantly, we really do need age cohorts, and maybe gender cohorts depending on the country and context that we're looking at; maybe some other distinctions. I should say gender distinctions among the age cohorts. And we may want to have other factors that vary by time.

But up to this point with the cohorts, we're still in an analytic solution world because we've just expanded the number of variables. We haven't actually made coefficients time-dependent.

So it's just good to know that you have an analytic solution ready for you. You can look at equilibrium, see where you're headed, even if you want to put in some changes over time.

I think that's all I wanted to say about the continuous versus discrete time. And I just want to quickly talk about Markov chain models, which use individual transition probabilities per period rather than transition rates or proportions. Does that matter?

First, the Markov property, just to

remind you, basically states that where you're

going, the future state is independent of past

states, given the present state. You have to say

given the present state or it's not true. But with

that condition, this is a very helpful assumption

to simplify models. And I hope Dr. Markov there

would agree with my explanation.

Then, of course, we have to scale up

these results for individuals, which hop from state

to state until eventually they hop to the death

state, and that's the end of that chain. We have

to scale that up to expected results over a

population.

I'm arguing here that Markov chain models

are fully consistent with the models above because

we can interpret the probabilities as the

proportions in a discrete time model. So if we

have a continuous time model, we can make it a

discrete time model, as I discussed.

Now instead of transition rates, k, we

have transition probabilities or survival -- I

should say transition proportions or survival

proportions, which we can then interpret as probabilities.

One thing to be careful about is that we don't confuse quantities that are zero to 1. So ask yourself, is this a zero to 1 quantity logically with quantities that are zero to infinity, which are rates?

My advice to you when you're trying to convert is to try minus the log of either the proportion or 1 minus the proportion, depending on how you define the proportion. And that'll get you started.

We can use the Markov transition probabilities to calculate prediction intervals on the model output such as prevalences. Now, if we're just trying to account for sample size effects for looking at a population of 10 or a million or a whole country, that really won't be that important by the time you get into the millions because the uncertainty of the mean due to sample size falls off with the squared event, and event is pretty big; that's going to get pretty

small.

But there are uncertainties in the probabilities themselves, which you can then model through Markov chain, Monte Carlo, and other simulation techniques, which are computationally expensive but are very helpful to understand the effects of uncertainty.

Finally, Markov chain and transition rate modeling can be combined where we use the Markov chain to tell us what transition happens next, and then we use the transition rate modeling, perhaps, with distributions on time to transition to tell us how long it takes, given the transition.

So in summary, continuous time is the natural view unless you believe in chronons, which are 10 to the minus 24 seconds or something -- I'd call that continuous -- units of time. But models can take discrete time steps with no loss of accuracy, and they're simpler to simulate. You can just step through time.

Either way, you can express simple models, at least, in terms of -- terms exponential

in time. And the point is that there is an equilibrium out there, though it may take a very long time to get there.

The Markov chain models don't use rates or proportions, but use probabilities to hop from state to state. But they're fully consistent with the deterministic models, as I showed.

So in conclusion, no, I don't think it matters whether you use continuous or discrete time, nor whether you are using Markov chain assumptions as you simulate your population, individuals in your population, or just use a deterministic rate calculation.

I'd like to acknowledge the support on a larger research project of JTI and the team, including Dr. Monica Lee of JTI and Rudy Gunawan, my Pharsight colleague. Thank you very much.

(Applause.)

Discussion - Panelists

MR. PAREDES: Wonderful. So we have 10 minutes. Why don't we have the speakers in the panelists. Come to the table a minute. We'll have

the panelists, the additional panelists, introduce

themselves and provide any comments, if they have

any, they would like to add to the discussion.

MS. MUHAMMAD-KAH: I'll start off. I am

Raheema Muhammad-Kah from Altria Client Services.

I am a research scientist in the modeling and

simulation group. And I am a biostatistician by

trade. I've worked on various clinical,

nonclinical, epidemiology, and behavioral studies

both in a public health and industrial setting.

Also, I've been working the past the two or three

yours on computational modeling.

So with that, I'd like to thank the CTP

for the opportunity to be part of this important

workshop. I'm going to be really brief because I

know we're running out of time. But the speakers

had very good presentations, and it was good to see

different aspects from different approaches in

terms of MRTP application and looking at health

policies; and then the great lesson by Dr. Poland.

But with that, looking at population

modeling in the multiple product landscape, these

are just some of the considerations. One of them

is it becomes very important for us to dedicate our

efforts at the very beginning in terms of clearly

defining our assumptions and our framework to

ensure that we don't over-complicate the model, and

that there is enough complexity so that decisions

can be made.

Secondly, choosing the right modeling

techniques. So we know that there's a vast span of

modeling techniques, from decision trees all the

way to complex system dynamics. So understanding

the advantages and limitations of these models will

help us to define our outcome of interest.

Some of the factors to look into are, is

it a cohort-level modeling or individual-based?

Also, the time horizon, are these long-term effects

that we're looking at or short-term effects?

So most of these models in this multiple

product landscape are lending themselves to the

more complex models. And as we saw in Annette's

presentation, where she took a life table approach

and looked at the MCMC or the Monte Carlo Markov

chain approach to actually measure uncertainty, and
then looking at your compartmental modeling, gives
you that complexity, range of complexity, that we
have.

But then I think one of the more
important points that have been mentioned this
morning and also yesterday are looking at hybrid
models and how we can leverage the opportunity of
the positive aspects of these different types of
modeling techniques.

Lastly, given the topic of the lack of
availability of data, which we've heard today and
yesterday, we need to have robust sensitivity
analyses so that we can identify the parameters
that really matter, and then also looking at a lot
of uncertainty analyses to bracket these range of
values.

With that, we would do a lot of what-if
scenarios on these models so that they would help
us to identify data gaps. And then also, it would
inform our new data collection process. And as
models are reiterative in nature, therefore as new

data come about, we will be able to continue to test and to improve our models.

That's all I have. Thanks.

MR. PAREDES: Thank you.

DR. BOONE: I'm Edward Boone. I'm from Virginia Commonwealth University. I'm an associate professor of statistics. I'm a Bayesian statistician; I love uncertainty quantification type topics. A lot of my research is focused on quantifying uncertainty. And I'd like to applaud the two modelers here, and thank the gentlemen for speaking about modeling in general.

Dr. Banchand's model, she actually attempts to put in uncertainty via probabilities, which Bayesians love. And Dr. Vugrin looked at uncertainty from a different perspective just by looking at, okay, what are some possible extremes and how does that impact our results and our models?

This is really important in this landscape of this uncertainty quantify in multiple products because new products are introduced, and

we don't know what these quantities are. Right?

So what we need to do is make sure that
when we're going through and doing these things,
that we try to specify the uncertainty going in,
like Dr. Banchand's model. That's incredibly
important. But we also have to realize that we can
take some of the information that we get from
models like Dr. Mendez's model and Dr. Levy's model
that can serve as, for example, prior distributions
for many of these parameters that go into the
secondary-level models. So you can use these
things to tie together these uncertainties.

So what I'm hoping is that not only
should we thank these modelers for doing these
sorts of things, but building the connections
between these models is going to be really, really
important, the policy-level models as well as these
population-level models that are interested in
people in their various compartments versus the
implication of policies.

Because ultimately, what needs to occur
is these parameters that everybody has in their

models are static. Right? And we're going to say

they're uncertain. But how many people in here

actually believe those parameters will stay static

into the future?

Everybody in here is probably interested

in changing policy, which will change the

parameters. Okay? So we should introduce also

this idea of forcing. In mathematical modeling, if

you have parameters that are going to be changing,

you force the parameter to change in various

directions as time goes on.

I know many models that keep things as

static, you're looking for the steady state. But

putting this forcing in here, you're pretty much

guaranteeing you're never going to hit steady state

because you're going to be changing the policies.

So my biggest thought is, uncertainty

quantification is huge in this because we don't

know what's going to happen. And we need to be

flexible enough to think about putting forcing in

these models because we're changing the policy as

we go along. We're going to be changing the

landscape as the product evolves.

So saying, oh, this is what's going to happen in 50 years, anybody thinks that that's true, I think, needs to take a step back and we need to say, this is what will happen with this amount of uncertainty or certainty associated with it. Thank you.

MR. PAREDES: All right. I think that we would like to get you to lunch because we have some --

DR. DRESLER: One of the suggestions we had -- and Dr. Flaherty, are you -- is it all right with you and Dr. Phillips if we do your session after lunch? Or if everybody just really wants to stay in the room for another half an hour or so before going to lunch?

The suggestion was that we hold that session after lunch, if that's okay with the two centers and okay with everybody else, which means you have a few more minutes left for questions.

MR. PAREDES: Okay. Any questions from the audience?

(No response.)

MR. PAREDES: Any comments from the
speakers about the previous comments?

DR. VUGRIN: At the risk of putting off
lunch for just a minute, I think one of the things
that Dr. Mendez said that was particularly
eloquently put earlier was, we think of uncertainty
as important, and I absolutely agree with that.
And we want to get the data, and we want to get
that fast. But sometimes there are other ways to
think about the problem. Look for the tipping
point. How bad do things have to be before the
benefits are undermined and done. And then
rationalizing and thinking through, can we hit that
tipping point or are we not going to.

So uncertainty analysis is important.
Couldn't agree more. We want to get the data.
We're never going to have it fast enough. So just
logical ways of thinking through the problem can
always supplement it, and knowing it's an
assumption. But there are a variety of techniques
that we can use to try to still answer these

critical questions.

MR. PAREDES: Wonderful. Thank you very much.

(Applause.)

DR. DRESLER: So let's be back and start at 1:00.

(Whereupon, at 11:58 a.m., a luncheon recess was taken.)

A F T E R N O O N   S E S S I O N

(12:59 p.m.)

Moderator – Carolyn Dresler

DR. DRESLER: Welcome back for the afternoon on our second day. And thank you for your patience in letting us move this session to right after lunch.

I realize that I had never introduced myself. I am Carolyn Dresler, and I'm the associate director for medical and health sciences in the Office of Science here at CTP, and very much work across the disciplines. So I'm very excited always to be involved with the workshops and TPSAC, et cetera.

This session that we have -- as you all know, you needed to register in order to attend the conference or in order to speak. So we had two presenters that are the next two people who we felt that we wanted to hear from, and we really didn't have a session for it.

So this is that session that we wanted to hear from these two speakers. The first one is

Brian Flaherty from the University of Washington. And he's been working in latent class and literature models for quite a long period of time, and has been working in those areas additionally in tobacco control. So it should be an interesting area for presentation.

Brian?

Presentation – Brian Flaherty

DR. FLAHERTY: Thank you very much for having me and for coming back after lunch. Lunch might be a good break because this is a rather different kind of talk than much of what we've been hearing, although I'd be quite game to talk to people about connecting this, perhaps, some of my stuff to some of the systems modeling.

I want to give a little more background after listening to the talks from the past couple days than what I actually have in my slides. This is empirical, so it's not mathematical modeling. I'm doing data analysis, multivariate data analysis.

The models are kind of similar to the

topic modeling presentations from yesterday

morning, the two classification talks that we

heard. And one of the things I like, we've talked

a lot about uncertainty and taking uncertainty into

account and presenting uncertainty.

Well, one of the things I'm focused on

here is how we think about uncertainty in these

models. And so I'm going to start with a

conceptual exercise. Hopefully, you won't find it

torturous, and how we think about uncertainty and

knowledge in this domain.

Just to give you an orientation to the

kind of stuff I do, so Gary Giovino yesterday gave

a talk about agents/hosts. So you might think of

my work as saying, how do we classify hosts? How

do we measure people? How do we organize them,

maybe parsimoniously, and integrate them in other

models?

As it says here, I'm using latent class

and mixture models. These are classification

models, statistical models for classification. And

then a common motivation for using these models is

that perhaps people in different subgroups respond

differently to a regulation, to an intervention, to

a treatment. And so that's a reason to look into

these models or study these models.

The presentation I have to give is

motivated by my own experience using these models

for 10 or 12 years, maybe more, but also how I see

people using them in the literature. And I'm going

to be balancing, talking about parsimony,

complexity, and knowledge, and trying to balance

those when we think about evaluating models.

So that's a background. I don't think

there's questions for this session. So if anybody

does want to talk to me, let me know during a break

or afterward.

So what I said, I'm going to try to do a

thought experiment first to walk you through and

contrast some things, factors that we use in

evaluating statistical models. And then I want to

bring that into the real-world context where it

gets worse, and then a brief empirical

illustration. And we'll see how far I get, and

might just jump to the punch line of the empirical

illustration.

What I want us to think about in this

little experiment are how these factors bear upon

statistical decisions, so thinking about the

knowledge we have, about the domain we're working

in, thinking about the preferences the researcher

is bring into the study. There's plenty of

criteria to think about for evaluating good

science, but I want to contrast or focus on, today,

parsimony and conservative testing.

Parsimony we've heard about a couple

times, preferring smaller models to larger models,

more complex models, all other things being equal.

And then this other idea of conservative testing

follows from an experimental world, where we like

to make things hard for the researcher to get what

he or she wants.

So I'm going to be really contrasting

those two. And I'm contrasting those because they

are not playing out well in how these models are

being used in the broader literature. So I'm going

to contrast them in what I'm hoping everyone here

is familiar with, is null hypothesis testing

paradigm, which I'm assuming everyone knows, and

I'd like to breeze through this quickly; kind of an

experimental paradigm. I've got a new treatment.

I want to see if it's better than standard care or

some control. We're focusing often on single

parameters and mean difference; treatment effect.

But often -- and this is where I want

people to be thinking about knowledge and not

having complete knowledge and statistical

decisions -- the researcher is usually coming into

this domain with a strong expectation that, well,

my treatment should be effective. If I know my

science, I've done a good job designing a good

treatment, and I expect to see a difference. I

expect to reject my null hypothesis, right, my

straw argument that there's no difference. I never

believed that, almost, in a null hypothesis

paradigm.

So I'm going to hope to reject this

thing. And my preference is to be able to say,

well, the results are so unusual, it's not
plausible to say there's no difference in the
population. It's not really plausible that my
treatment doesn't have an effect. So I'm going to
conclude that my treatment has an effect. Right?
So that's kind of the framework of null hypothesis
testing.

Model fitting is where latent class and
mixture models come in, path analysis, structural
equation models, factor analysis. All kinds of the
more complex multivariate models are in this world
of model fitting.

Instead of single parameter tests, we
often care about multi-parameter tests. And this
is an important point, and I think being ignored in
the current literature. The researcher here
hypothesizes a structure, an association structure,
their data based on their understanding of the
phenomena. And that is the null hypothesis model.

So the null, the model that I care about
in this framework of model fitting, is the null.
And so if I have a good understanding of my domain,

I should be able to propose a pretty accurate model of the phenomena and then retain it. That's my hope. And the alternative here is some more complex model.

So what I wanted to do -- the rows here are different factors, my knowledge, my understanding of my science, the lack thereof, the converse, the researcher's preferences, and then these two factors, conservative testing and parsimony.

There are all kinds of things that drive study results. Right? So right now, I'd like us all to imagine we've done the study perfectly well, and we also don't have any weird data set. We don't have some weird data set on the tail We have a nice, representative data set, and we've done the study to the best of our abilities. We've designed it right. We've got the best measures. We've got everything, good samples.

So if I have a lot of knowledge about my domain, and null hypothesis testing, as I said, I'm going to reject it, if I have a good knowledge of

my domain and model fitting, I'm going to keep it.

I'm going to retain in. Lack thereof, I'm going to

retain the null disappointedly, and model fitting,

I'm going to reject it to the degree I don't have a

good understanding of my domain.

The researcher preference, as we said,

reject the null here in null hypothesis testing,

and retain the null in model fitting. And then I

just put in parentheses underneath, by rejecting

the null, we're concluding complexity. We're

concluding that things are more complicated;

whereas by retaining the null in a model fitting

framework, we're concluding simplicity. We're

sticking with parsimony here.

So now if we jump to this idea of

conservative testing, in null hypothesis testing,

the idea of being conservative pushes us to keep

our null. If we want to be conservative, we're

going to make it harder to reject the null in null

hypothesis testing.

In model fitting, if we want to be

conservative, we're going to make it easier to

reject the null. We want to make it easier to

throw out what I'm putting forward as

the association structure explaining the data.

Parsimony here, if we're interested in parsimony,

that's going to push us to retain the null here,

and it's going to push us to retain the null here.

So what I don't think is really being

appreciated now is this idea that in model fitting,

parsimony is pushing in the same direction as the

researcher preference. It's pushing us to keep

models that the researcher's putting forward as my

model of interest, my model of theoretical

interest.

The reason this is important is -- well,

let me come back to this in a minute -- is in the

real world we've got a ton of uncertainty in

evaluating models. And people are falling to

parsimony because it's easy. You can say, oh,

simpler is better.

So let me now step to this one and give a

very brief sense of these models. So latent class

and mixture models, they're being used all over the

place. They're growing quite a lot across domains
of science.

They're proposing qualitatively different
subgroups. So we've talked about poly tobacco use
or nicotine use. We've talked about patterns of
smokers. We've talked about transitions among
states. These models map onto those phenomena.
However, the latent class and mixtures, we're
inferring. They're not directly measured. They're
inferred classes or states. And I motivated it
already. I said that already.

Often, though, in these models -- and
this is also important -- is a very little
theory -- very often we don't have the right amount
of theory or enough theory to postulate what the
classes are or what the structure should be. So
very often it's exploratory. So I mention this.

When we get into the real world and are
fitting them to real data sets, it's often a great
deal of uncertainty on how these models are being
used. And so we follow parsimony, and let's just
jump to the punch line here. And then if I have a

couple moments, I'll show a very brief example.

But when we don't know what's going on very much in these models, we don't have strong expectations or strong theory, and we have uncertainty in our model selection, I think when we have uncertainty across a set of models, we should be leaning towards a little more complex models rather than the smallest models because the small models overstate how well we understand things, and the more complex models are better quantifying the uncertainty, I think, about the behavior.

So I have come to this -- I've been working latent class and mixture models. I got into tobacco because people talk about different smoking patterns. And when we talk about different patterns and combinations of product use and stuff, I think those phenomena could map onto here as well.

But the example I'm going to talk about is brief -- different smoking patterns, and there's very little theory, very little guidance about what

those smoking groups might be. And people

disagree. They cut it off differently. They form

the groups differently. And I'm taking an

empirical approach.

This is just one of the analyses I've

done. This is 26 years old from a 2004 National

Survey on Drug Use and Health. I'm only looking at

a white male subset, 26 and up, so keep that in

mind. And I'm only looking at people reporting

smoking, so about 2,000 people.

I'm not going to spend too much time

talking about it. It's a very small model. It's

almost as small as it gets. And 2,000 people, this

is almost an ideal situation for using these kinds

of models, and I'm still running into problems

making model selection and doing good inference.

So these are the data. There's two items

on history and two items on -- I'm sorry, two items

on history and two items on recent use. And the

punch line for this group is more, so there's

uncertainty in model selection.

These are models with different numbers

of classes, two, three, four, five classes. These

are fit statistics, so in a Pearson's x squared and

a 2 by 2 table, independence, same kind of idea.

And then AIC and BIC, these are two

parsimony -- trying to balance parsimony

complexity.

This table shows up every time I do an

analysis, practically. BIC always says, fewer

classes. So if I go with BIC, I'm emphasizing

parsimony over fit. And this is what everybody

does in the literature.

You go read a paper, and you'll see a

table like this, and I'll say, well, three classes

and four classes both fit. But our BIC says three

classes; that's what I'm going to go with.

I'm not going to bother interpreting the

four-class model -- I mean, the three-class

model -- because of time, but I just want to draw

out that in the four-class model, the one that if

we were falling to a little more complexity, we

would be talking about, the fourth class is a group

of new smokers. These are people 26 and up. Two

percent; not a huge percent, but 2 percent in the

U.S. population is big. And 26 and up is not what

we expect for new smokers.

So if you just used BIC, this wouldn't

show up. In my discussion and my results, I

wouldn't bother talking about them. Leaning

towards the more complex model, you see more

patterns or more rich data. This is a summary of

what's different between the two. And let me just

jump to the end because I'm a minute over.

When we talk about balancing uncertainty

or quantifying uncertainty, I'm trying to keep the

uncertainty in the model selection and think about

how model selection interacts with how it portrays

what we understand about the world or about the

phenomena. Yes. I'll just stop there. Thank you.

(Applause.)

DR. DRESLER: That is an opposite message

than what we've been hearing, isn't it? Go for

more complex and more simple.

Our next speaker is Dr. Carl Phillips,

who is from CASAA, which is Consumer Advocates for

Smokefree Alternatives Association. And he has

been working for many years in the area of modeling

and in advocacy in the tobacco-related area.

Dr. Phillips?

Presentation – Carl Phillips

DR. PHILLIPS: The thing about doing a

presentation, a conceptual presentation, late in

the day in a workshop is a lot of what you've done

has already been said, although that's probably not

nearly as annoying as having someone come up late

in the day and say, there's been a fatal omission

from everything we've talked about so far.

The argument that I want to make here is

that existing models of tobacco use basically only

either calculate hypotheticals or extend past

trends, or I should have added to that

econometrically explain past trends. But this is

not useful for understanding today's tobacco use,

or in particular, for regulatory requirements.

Instead, we need predictive models, which

require an inclusion of mechanisms, the "why" of

the system. And the mechanism that drives tobacco

use, the primary mechanism, is welfare economics,

which is to say consumer preferences and choice.

There are no simple constants in the

world of tobacco use. Now, if you're trying to

model something like the flow of water through a

cooling system, you can simplify out the mechanisms

because every 45-degree bend in a pipe has the same

effect on water pressure no matter what year it is,

no matter what country you're in.

But tobacco use behavior doesn't resemble

hydraulics. Now, you can be excused for thinking

that it did if you looked at some of the models

that exist. But in fairness to those, decades of

near sameness did create an illusion of constancy;

looking at a single product and a relatively

unchanging pattern of population behaviors make it

look that way.

But the reality is, there are no physical

constants in the social sciences, ever.

Epidemiology 101 teaches in the first week, or it

ought to, that people, place and time matter, that

populations vary and change, and thus no

epidemiologic estimate is ever anywhere close to
being a constant. And that doesn't even take into
consideration changing product options in the world
of tobacco.

Moreover, we are trying now to make
predictions that are out of sample, which is to
say predicting impacts of novel products and new
educational efforts. And furthermore, it's in a
world already characterized by massive novelty,
even apart from any individual regulatory decision.

Therefore, it's impossible to just
catalogue what's come before and figure out where
within the bounds of what we've seen before
something new fits, and just plug it in there. It
just doesn't cover the right space. Nor is it
sufficient for many of the purposes we have in mind
to just look at the hypotheticals about, if x, then
y.

On the other hand, I would argue that
hypotheticals' best educated guesses about inputs
are better than naive extrapolations of previously
high-level epidemiologic results that, say, look

at a different product in a different population,

and most importantly, under radically different

circumstances, and assume that this can predict a

novel situation.

So to take a cartoon example that's not

too far from the real world, assume that a novel

THR product, tobacco harm reduction product, in

2014 will perform the same as snus did in 2004. Or

assume that smoking patterns in the era of THR

resemble those from decades past.

I would argue that such use of high-level

result is not science. I would call it

superstition. That is, we saw x follow y once in

the past, so we're going to conclude that that's

what always happens. We have no idea why that

might be the case, but we saw it once and so we're

going to continue to believe that it's true.

The analogy I like to use is imagine

predicting the trajectory of a rocket launch from

Korea using the prevalent methods from

epidemiology, public health, and tobacco behavior

modeling. You'd predict that it would be off the

coast of Florida a couple of minutes after it
launched because that's where most satellite
rockets are. That's what the data shows.

Now, of course, we can predict where the
rocket is really going to be based on our knowledge
of mechanics. But if we didn't already have that
knowledge of mechanics, we could go back to the
Florida launch data and figure out how rockets
work, and then apply that to the new situation by
paying attention to the mechanism. But we don't
get there if we just look at where the average
rocket is two minutes after it launches.

So mechanisms are what cause the moving
parts in a model to do what they do. And they,
unlike high-level results, are relatively
consistent. Now, in our case the mechanisms are
consumer preferences, individual tradeoffs among
different costs and benefits, people's knowledge,
social factors, supply factors, and so forth.

Now, a few starting points that are
necessary to understand if we're to have any hope
of realistically modeling consumer behavior.

People like to use tobacco products. People choose
to use tobacco products.

         Many users switch products when they
perceive that an alternative provides better net
benefits, such as the much lower risk is appealing
even if the product they perhaps don't like quite
as much as smoking.

         More generally, tobacco consumers are
trying to do what people are always doing when
they're making consumption decisions. They're not
doing it perfectly. They're not doing it perfectly
rationally. But they are trying to maximize their
welfare.

         So the basic economics of tobacco product
choice, what has to be at the heart of any model
that actually predicts behavior, that tries to
figure out what's going on out of sample, has to be
something like this.

         People have a benefit from each
particular product, which is primarily driven by
product quality, which is basically how good the
product is, how much they like it, but also has

factors such as product knowledge as well as the

acceptance and popularity of the product. In some

sense, the Sandia model you saw yesterday is all

about it just being acceptance and popularity.

That's critical, but it's certainly not everything,

and I would say it's definitely secondary to

product quality.

Then there are purchase prices. There

are the perceived health effects, which are a

combination of the real health effects and people's

level of education or disinformation about the true

value. And then each individual has a tradeoff

among these, and other considerations if you want,

but I think this is the core of it.

The tradeoff leads to preferences, how

much value they get out of using a particular

product versus none at all, and preferences lead to

decisions. And you can add something into the dot

dot dot box about how preferences lead to

decisions, depending on how far you want to get

into behavioral economics.

One thing that is absolutely critical no

matter how far you're getting in is you have to

have some kind of hysteresis effect, that people

tend to continue doing whatever it is that they're

doing. And that's not just in one direction.

That's in any direction from any possible change.

So actors each make their own decision

based on something that looks very much like this.

And I have to say that this largely forces us into

an ABM method-type world if for no other reason

than because of the fundamental heterogeneity,

which has huge implications for this. Without this

mechanism built into the model, it tends to be more

of a calculator of hypotheticals, and it certainly

doesn't let you go out of sample.

Of course, effective models have to be

able to explain the empirical observations that

we've seen to date. The model that you create that

is based on these mechanisms needs to be able to

match decreases in smoking in the late 20th

century, recent adoption of THR products in various

markets, and so forth. But the key is that it

tells us far more than a simple review of past high

level outcomes can once you make a mechanistic

explanation for what it is that we've seen before.

The parameters in that flow chart are

definitely still not constant. As I said, there

are no constants. But they're a lot closer to

being constant than any particular estimate about

what happened under particular circumstances in the

past, and therefore, they transfer to analyses of

novel situations, similar to being able to figure

out where a rocket is going no matter where it's

launched from.

This allows us to answer questions like

how much effect might mass communication about a

product's low risk have, does a product quality

change have much impact, or is the social buzz more

important, and so forth.

Why? Why do we need this? Well, by my

reading, the MRTP guidance -- and now apparently

some of the substantial equivalence applications

also -- call for predictions about what people will

do under novel circumstances. Of course, even

apart from regulation, an awful lot of us would

really like to make predictions and figure out the

key levers about what is going to encourage product

switching and so forth.

I would argue, furthermore, that even

hypothetical calculators, like how many people are

going to die over the 21st century, are

increasingly useless because they fail to deal with

novelties. And any calculation about what the

future looks like that's more or less the same

today as it was in 2009 before e-cigarettes took

off has to be doing something very wrong. But how

can you figure out how to incorporate how e-

cigarettes are behaving? Well, you need to drill

down to the actual mechanisms of the situation.

Now, wait, someone might ask. Isn't

preference-based modeling far too complicated,

and doesn't it produce results that are far too

uncertain? Well, yes. It's complicated, and the

results definitely are uncertain. But I will argue

that realistic modeling has a good chance of giving

correct answers, whereas seemingly precise but

completely hypothetical and unrealistic modeling

does not.

Furthermore, as a couple of speakers have pointed out -- Dr. Levy and Dr. Flaherty in particular -- you actually can gain a lot with this type of complexity. You eliminate some other areas of complexity that are far too assumption-dependent by doing this.

The argument sometimes gets made that this departs from solid science because it's not basing the parameters on what we observed happening according to some study that was published in a journal. But as I pointed out, I think I would call using numbers like that superstition. This is the solid science of the behavior.

Furthermore, those who claim to be just looking at the facts are actually engaging in thought-free theorizing. All data is interpreted through theory; it's just a question of what theory you start with.

So you observe previous behavior, and you might say, assume the transition probabilities for the same actions say the same, the scare quotes

suggesting that, of course, no action is the same in an ever-changing world. So you get a precise-seeming calculation that bears little resemblance to the out-of-sample future that we are trying to figure out.

Alternatively, if you drill down to the mechanisms, that people act based on preferences and so forth, it's more difficult, but you can make predictions about previously unobserved but characterizable situations.

Now, this isn't to imply that people make perfectly rational choices about tobacco use or anything. Obviously, they don't. I built in major irrationalities to adjust that flow chart sketch of people's decisions already, and more can certainly be added if there's empirical support for adding them.

But my argument is that starting with economics, rationality, volition, people trying to be happy, is much better than assuming that tobacco use is caused by demonic possession, which in my mind describes the tobacco control notion of the

underlying notions.

Or if you prefer a light mythos to a dark mythos, you can call it the theory of immaculate causation; that is, claims about why a particular change will have a particular effect are usually a magic asterisk, where there's no explanation for why that might be the case.

As you might expect from that, most of those predictions are usually wrong, with the obvious exception being that the responses to price increases are usually pretty solid. But notice that that's a case where there is a concession that this isn't about demonic possession. This is about maximizing utility, and people are responding to things in a rational, economic way.

One last point on this. It might seem that this complicated mechanism-based modeling is not needed once you have surveillance data because then you have the data. But mostly all we can do with surveillance data if we're lacking a realistic model is just extend the trend lines.

So, for example, we have data about

rapidly accelerating uptake of e-cigarettes, but
what's the shape of the curve? Where will it level
off? How will introduction of new competing THR
products change the future of e-cigarettes over the
next few years?

These are questions that you'll have a
shot at answering if you've actually drilled down
to the underlying mechanics. But a few years of
surveillance do not give you much of a read on
them.

Just to take the opportunity to mention a
few results from the modeling that I've done, both
static and agent-based modeling, that starts with
these premises. And I obviously don't have time to
talk about any of the details about how I got
there.

But one thing that immediately comes out
of any use of this underlying mechanistic paradigm
is that availability or knowledge about low-risk
products will increase tobacco use ceteris paribus.
It will do this because it increases the net
benefits of using a particular product, as compared

to cigarettes, quite dramatically because it lowers the cost. This is something that regulation and all policy-making is going to have to come to grips with. Tobacco will be more popular when it's low risk.

The hysteresis and social knowledge and acceptability effects are huge. You heard a whole talk on that, so I won't go into it. Right now everything is about the disequilibrium; that is, figuring out what's going to happen in equilibrium is not very interesting. Not only in the long run are we all dead, but in the next year or two, everything's going to change yet again. So trying to figure out what's going on with the disequilibrium is critical.

Minor variations in product quality can make huge differences in outcome. So e-cigarettes exploded; Camel Sticks flopped. Most users who have really given them both a good, solid try think that they're both fairly high quality alternatives to cigarettes.

Yet what happened? If we can't explain

that, we have a problem, and existing systems

models don't tend to predict either of those

outcomes, but rather predict a slow, steady growth

with a little bit of uncertainty fuzz around the

edges of it. There's necessarily huge uncertainty

in any estimates. But as other people have talked

about, it's possible to use these models to assess

where the uncertainty is coming from and deal with

it.

Finally, I want to thank BAT, JTI, and

CASAA for support of the larger project that this

is part of. But this is all my own opinions and my

own work, and in fact, this I had worked out

entirely before I had any support for the project.

I know there's no question time there.

My email address is at the bottom. Thank you.

(Applause.)

Moderator – Carolyn Dresler

DR. DRESLER: Thank you very much.

The next session, next to last session,

is Population Models to Account for Quality of Life

Measures. We're going to have a little bit change

in the order, but we'll still have -- the first speaker is Dr. Marie Ng, assistant professor of health at the Institute for Health Metrics and Evaluation at the University of Washington.

She's been working in modeling and quantitative methods for quite a period of time in several places around the world, and so she will be presenting today. Her presentation is on measuring global burden of smoking methods and challenge.

Dr. Ng?

Presentation – Marie Ng

DR. NG: It's a pleasure to be here today to share with you the methods and challenges for measuring global burden of smoking. My presentation is going to focus on some of the research work conducted at our institute for estimating disease burden that's attributable to smoking.

I'm going to start off by talking about a metric that we use to measure disease burden. Specifically, it's called DALY, disability-adjusted life years. I'm going to describe what it is, why

we use it, and how it's being calculated.

Then I'm going to talk about one of the applications of DALY in a study called Global Burden of Disease. It is led by our institute director, Dr. Christopher Murray. It's a very interesting study. It measures the distribution of disease, injuries, and risk factors worldwide. I'm going to focus specifically on how in that study we try to attribute the burden of disease that is related to smoking, and some of the challenges involved.

So what is DALY? DALY is a summary metric of population health. It is a measure that captures health loss that's related to both death and illness. So essentially, it measures both mortality and morbidity. Another feature of DALY is that rather than capturing health expectancy, it tries to capture health gap. It represents the state of population health compared to an ideal state.

Now, why do we use DALY? DALY can be calculated for a specific disease; therefore, it is

very useful when you're trying to measure burden of

disease worldwide. Another feature, as I mentioned

just now, is that DALY is measuring health gap. So

it's comparing to an ideal state. The egalitarian

principle built into this metric is one of the

appealing features of the use. Also, we can

compare countries and compare across time using

this metric.

So how do we calculate DALY? It's not

that complicated. There are only two components in

the DALY metric. One is YLL, years of life lost

due to premature mortality. The second component

is YLD, years lived with disability.

If I can simplify this concept, let me

use a diagram here. Imagine that this rectangle

here represents the ideal achievable health. So on

the Y axis here, you see is the health state, and

then on the X axis here, it represents the idea

lifespan one can achieve. So the area under the

rectangle is the total unit of healthy life that

ideally is achievable.

Now, suppose, unfortunately, you die at

the age of 60. Now, at the age of 60 we know that the standard life expectancy at the age of 60 is 87.8. So if you die before that, it means that you have lost 29.8 years. So your years lost due to premature mortality is 27.8.

Now, suppose that five years before that, you have some condition, and that condition discounts your health status, and the disability weight related to that condition is .3 here. So your years of life lost due to this disability is five years times the disability weight. The YLD here is 1.5.

Suppose, so unfortunately, three years before that you have another condition, and that condition again discounts your total health status. So three years, this time the disability weight associated with that condition is .2. So your YLD in this case is .6.

Now, DALY simply is the sum of all these three. So your DALY in this case was 29.9. Now, I should emphasize that what I have just presented is definitely an oversimplification of the concept of

DALY. But this is just to show you what DALY is
made up of and the idea behind it.

DALY as a population metric is calculated
generally at the population level, and this is how
it is calculated. The YLL is the number of deaths
at age A at a population level, and it's multiplied
by the extended life expectancy at that age. So
this is how at the population level we compute YLL.

As for YLD, it's disease-specific. So
it's PI, which is the prevalence of a disease
multiplied by the disability weight, ranged from
zero, full health, to 1, death.

Now, before moving forward, I just want
to quickly contrast DALY with another well-known
metric known as QALY, quality-adjusted life years,
which some people are more familiar with.

So both metric measures are a combined
measure of mortality and morbidity. However, there
are some differences between the two. First of
all, the origins of the two metrics are slightly
different. DALY is coming from health metric,
whereas QALY is from health economics.

Because of the difference in origin, you see that the application of them also differ. You will see DALY applied in most studies related to disease burden, whereas QALY is often used in cost-effectiveness analysis. However, that's not absolute. You will still find studies in cost-effectiveness analysis using DALY, and you will see disease burden studies use QALY.

Another also very important difference between DALY and QALY is DALY measures the years in perfect health lost, whereas QALY measures the years in perfect health gained. So one is lost and one is gained. So ideally, what you want to do is you want to minimize QALY and maximize DALY. So this is one of the major differences.

Both QALY and DALY utilize some weighting. For the disability weight used in DALY, it is generally disease-specific; versus in QALY, the disability weight is related to the general health status. And it's also important to note that for DALY, the disability weight, the zero refers to full health; 1 refers to death. It is

the opposite for QALY. So that's why 1 is
measuring loss and the other is measuring gain.

Now, the differences that I point out
here are really, really general differences. You
will see there are a lot of discussion papers
comparing and contrasting the two, and some
researchers definitely have a strong preference and
opinion for one over the other. But I'm not going
to get into the debate here.

Now, one of the very fascinating uses of
DALY is in the Global Burden of Disease study. In
fact, this is where DALY was originated. The
Global Burden of Disease study was started in 1990,
and the most recent round is the GBD 2010. GBD
2010 was funded by the Gates Foundation and is led
by our research institute director, Chris Murray.

GBD is a systemic, scientific
investigation trying to measure the health status
of 187 countries from 1990 to 2010, for GBD 2010.
It captures the burden of disease related to 291
conditions, over 1,000 sequelae, and 67 risk
factors. And among the 67 risk factors, tobacco

smoking is one of them. So I'm going to focus

mostly on describing mostly how we attribute the

disease burden related to smoking here.

So what did we find in GBD 2010? In GBD

2010, we found that tobacco is a major driver in

global health patterns. It contributed to

5.7 million deaths, 6.9 percent of global YLLs, and

5.5 percent of global DALYs.

Now let me just go in depth a little bit

to show you some of the results from GBD. So all

of the results from GBD is on our website in the

visualization tool. So if you click on GBD

Compare, what you see here is a tree map. This is

essentially just a square pie chart, and the blue

box here shows the DALY attributable to non-

communicable disease. The orange boxes here are

the DALYs contributed by communicable disease, and

the green boxes here shows the DALYs contributed by

injuries. Now, the larger the box, the heavier the

burden. And this is all sex, all age in 2010.

Now, if we scroll down through the years,

you see that the box changes. So this is the DALY

distribution in 1990. You can see the burden in
1990 was more heavily related to communicable
disease, whereas in 2010 it was more related to
non-communicable disease.

Now, if I change the chart to risk factor
attribution and look at tobacco smoking, you see
that there are some shaded areas in the box. The
shaded area here represents the DALY, the burden
that's attributable specifically to smoking. So
you can look at here we have lung cancer;
.96 percent of the DALY related to lung cancer was
attributed to smoking.

There are also many other interesting
features here. For example, you can look at the
stacked bar chart. This bar chart shows the
relative importance of the different risk factors
that's related to the burden of disease.

I'm going to pause here, and if you are
interested in looking into more, you can go to
our website or you can just simply Google "GBD
visualization," and you'll be able to find more.

Now, I have been talking about using the

term "attributable burden" without really defining it. So exactly what is attributable burden? By definition, attributable burden is the difference between burden currently observed and the burden that we would have observed if the population exposure followed a different pattern.

The calculation of the mathematics behind it is actually very, very simple. So AB here stands for attributable burden. PF is the population-attributable fraction of the risk that is associated with a particular disease J. And B sub J here is the burden associated with disease J. So it's just very simple products, multiplication and summation.

Now, as you can see here, population-attributable fraction is a very key component in this computation, and it is computed using this formula here. The R function here is the relative risk, P refers to the population distribution of exposure that we observed, and P prime X is the distribution of population exposure, a counterfactual distribution.

So remember that attributable burden is the difference of burden between the observed and the counterfactual. And it is actually presented here, in the numerator.

Just to present this in diagram form, suppose the pie here represents the total DALYs that's attributed to a disease, so the DALYs, for example, of lung cancer. And part of the DALYs that's related to lung cancer is contributed by smoking, some risk factor here.

Now, suppose no one smoked. We would have observed the total DALY to be 15,000. The difference between these two, which is 5,000, will be the DALYs attributed to smoking. So essentially, the risk factor-attributable DALYs is just the population-attributable fraction times the total DALY of the disease. But as you know, one risk factor can contribute to multiple diseases. So the total attributable burden for a risk factor is the summation of all the diseases.

Here's an example. Suppose this is the total DALY of disease A, and the risk factor that

contributed to disease A contributed 5 percent of
the attributable fraction. So the attributable
DALYs of risk factor to disease 1, 5 percent times
20,000, this would be 1,000.

The same risk factor also contributed to
another disease. In this case, it contributed
20 percent of another disease, which constituted
30,000 DALYs. As a result, the DALYs attributed to
the risk factor to the disease is 6,000. So the
total attributable burden of the risk factor in
this case is the sum between the two, which is
7,000. So this is the concept behind attributable
burden of a risk factor.

Now, estimating attributable burden is
very useful. Particularly, it enables us to
identify the major contributor of health loss, and
also allows us to compare across countries to see
how countries differ in terms of the impact of the
risk factors. However, it is a very difficult task
to do. Estimating attributable burden is not a
trivial task, and estimating population-
attributable fraction turns out also to be very

difficult.

If you remember the equation that I just presented, one of the major determinants of calculating population-attributable fraction is to know the population exposure distribution. In smoking, it is not easy to do at all.

The two typical metrics for measuring exposure in smoking, one is prevalence, which shows the penetration level of tobacco use; the other one is consumption, which is related to occurrence of diseases. However, as mentioned in the panel yesterday, none of these metrics are consistently monitored in the public health system.

So when we are doing this at a global level, the situation is even more dire. We are faced with a situation where we have very, very sparse data, inconsistency in the estimate of data sources, different definition of frequency and types, and we also have reports that will report at an aggregated level rather than at the age/sex-specific level that we want.

Just to show you some examples of what we

see at the global level, the graph here shows the

male smoking prevalence of Timor-Leste. The blue

points are the raw data. So as you can see, each

graph represents a specific age group. As you can

see, there's only one data point in many age

groups. When we are lucky at the age group of 15

to 19, we get two data points. And in many age

groups, we don't even have any data. So this is

one of the challenges as we are proceeding with our

analysis globally.

Even for countries with more data, here's

an example for China. Again, this is the male

smoking prevalence for China. Yes, we have a lot

more data for China; however, if you look this

graph here showing the male smoking prevalence for

age group 15 to 19, in 2010, we have two data

points from two different surveys. One is giving

us an estimate of smoking prevalence of 10 percent.

The other one is giving us an estimate of smoking

prevalence. So the inconsistency in the estimate

from different data sources is a real challenge to

the analysis.

As part of the GBD 2013, we did a study to estimate the global levels and trends in smoking preference and consumption of cigarettes in 187 countries by age, by sex, from 1980 to 2012.

To address the challenges that I just mentioned, we tried to aggregate or tried to acquire data from as many data sources as possible. We ended up getting data sources, unique data sources, of over 20,000 from national surveys, from international agencies.

The idea behind getting different sources of data is that through triangulation of different sources, we hope that we will get closer to approximating the truth. And to integrate this information, we applied data synthesis strategies such as Gaussian process regression.

Just to show you some of the results, here's the graph that I just showed in China. The blue line here is our estimate, so each graph represents a specific age group. So the blue line is our estimate. The shaded area is the uncertainty interval.

This set of graphs shows the male smoking prevalence by age group for Indonesia, and this set of graphs shows the smoking prevalence for the United States. As you can see the difference, in the United States we have a lot more data, although yesterday we were complaining that the data is still not enough. But compared to many countries, the United States does have a lot of data. And you can see the estimate actually follows the data very, very closely for countries with nice, consistent, abundant data.

This is the graph that I just showed on Timor-Leste. Very sadly, for many countries we don't have a lot of data, and as a result, the estimate is derived based on a lot of strength borrowing across time and space.

So just to summarize, I began this talk by introducing the concept of DALY, which is a useful metric for quantifying disease burden. One of the crucial components to estimate attributable burden of a risk is the exposure.

However, estimating the exposure of

smoking is definitely not an trivial task,

particularly where faced with a lot of data

inconsistency and data spareness issues. And to

maximize the use of existing data, some of the

modern technologies, statistical methods for

synthesizing data, could potentially be useful.

Thank you.

(Applause.)

DR. DRESLER: We're going to change the

order just a bit in order to accommodate some

flights. So we're going to change the order and

we're going to have Dr. John Ware come present.

In trying to figure out how to introduce

him, who has quite a very significant, very

impressive history, which you probably have heard

very long introductions, so I'll try and make it

short. But it's quite impressive.

He's professor and chief of outcomes

major division in the Department of Quantitative

Health Sciences at the University of Massachusetts

Medical School, and he's a member of the Institute

of Medicine, the National Academy of Sciences, and

he's published more than 400 peer-reviewed articles.

You can start there because that's very, very impressive there. But he also has numerous awards, and I'll just say one, the 2003 President's Award for International Quality of Life Research. So he's done a lot of work across broadly in this area, and we're very pleased to have him.

Dr. Ware?

Presentation - John Ware

DR. WARE: It's a real pleasure. Thank you for the kind comments. A real pleasure to be here.

The FDA is tremendously important and influential in the field, not just in regulatory decisions about quality of life benefits of drugs and devices. The whole industry looks to them for help with conceptualizing and measuring health. So it's a tremendous opportunity to be here today.

I'm a psychometrician, which is the theory and methods of measurement. So the models that I'm going to talk about very briefly are

measurement models. They're models of the

underlying assumptions that are behind the scoring

and the construction of psychometric tools.

I'll be talking about health-related

quality of life or patient-reported outcomes. And

to put it in the context of what we just heard,

these are assessments at the individual level that

you can aggregate to look at, for example, current,

former, and never-smokers.

Measurement begins with a conceptual

framework, so very quickly we're going to take a

pass through what is health. Of course, health is

one dimension of quality of life, and it's the

health-related part of quality of life that I'm

going to be talking about.

When I first saw the WHO constitution

definition of health, I saw two things that were

game-changers. Number one, it's a multidimensional

concept. So there are distinct components of

health, and we should be measuring them and

interpreting them separately. And we should

understand them well before we summarize them into

a single number.

The second thing that I picked up in this that's served me very well and is hugely important to what we're talking about in this meeting is that it has a range. And that range includes but is not limited to disease and infirmity, and it goes all the way up into well-being.

So we really want to make sure that we have a good match between where our people are and where our measures are in terms of the major dimensions of health. And this is a huge issue, and I'll conclude very quickly with five suggestions for going forward.

So how do we operationally define health? What I'm going to say right now, I think, really applies to every major component of health and every dimension of health. And at the core of this conceptual framework -- and I'm going to use the metaphor of an onion -- is your bodily structure. Is anything missing? And how well is it working?

One of my friends at IOM called that "organ functioning" as opposed to human

functioning, which is the rest of the layers that I'm going to talk about.

The next layer is the human experience of disease or some other condition or risk in specific symptoms. The next three, at the risk of oversimplifying the whole thing, is what health-related quality of life is all about, and it comes down to three things.

First of all, how does it feel? And that's the ill-being, the fatigue, the depression, the pain, and the well-being. Yesterday there was some talk about happiness. I would include happiness in well-being. Those are subjective things. But with good psychometrics, we can get reproducible scores for them.

The next big part of health-related quality of life is about what you're able to do. And the code we use for that is functioning, such as physical functioning. And that's the one measurement example I'm going to show you in the limited time that I have.

But the last point I want to make, no

matter what your score is on everything above in
this list, if it isn't good enough for you, that's
a major thing to know. So the cognitive evaluation
of your health state, regardless of the rest of the
evidence about it, is a huge predictor of just
about everything we want to predict in care. And
smoking hits that directly and substantially. I'll
come back to that in a minute.

So let's talk through this now in a
specific example. And I contrived this
specifically for the topic today. This conceptual
framework for outcomes, starting on the far left
with the most specific, the most objective
biomarkers -- and there's a huge, large number of
those biomarkers.

If I were going to make a plea to this
group that we need help in order to evaluate the
rest of the model, we can't have hundreds of
biomarkers. We need a few. We need them
aggregated. We need some order to that madness so
we can study the relationship of the change in
those biomarkers to the experience of the risk and

consequences of smoking and the benefits of

changing smoking behavior.

So what are the other three boxes? The

first are the symptoms that are specific to

smoking, such as a smoker's cough. The next, and

this is the beginning of health-related quality of

life, is the impact that is attributed to smoking.

Does it limit your everyday activities or your

enjoyment or your quality of life?

Finally, general or generic measures of

health that are not specific to any disease or

condition. So this is a continuum of specificity

to genericity in outcome assessment.

Now, this is the content of the field

over the last 40 years, on the far left, 1970, the

sickness impact profile. On the far right are the

utility-type measures that were just discussed by

the previous speaker. On the left of the utility

measures are basically the psychometric tools of

the 1970s, like the SIP, the 1980s, like the

Nottingham Health Profile, and the 1990s, like the

medical outcomes study and the MOS short form.

The new source of excellent tools for us to use in evaluating the quality of life implications of smoking are the patient-reported outcomes measurement information series tools.

With regard to all these tools, one very important point that I would make is our preoccupation with negative definitions of health. Defining a good year of life is a year without a disability, or defining physical functioning is the absence of a physical limitation, or vitality is the absence of fatigue.

Not only is this wrong in terms of capturing all of the human differences that are important that are affected by smoking or that are the side effects of medication, it creates a ceiling effect, an effect in which such a large proportion of the population has a perfect score, we can't see the benefits of changing a risk factor because their scores are already at the ceiling or the highest level for that particular measure.

Very quickly, according to the British Medical Journal, the SF-36 short form we developed

for the MOS is still the most widely used tool in clinical research. It has 36 items. It has eight domains. Those domains can be summarized into two component summaries.

Then, in the spirit of what we just heard by the previous speaker, a very important contribution, an award-winning contribution, was made by John Brazier and his colleagues at Sheffield. They took the psychometric description of a good year of life, and they gathered preference data so that it could be scored as a utility index where zero is death and 1 is a perfect year without any decrements in any of this descriptive system.

So what is the importance of that? As the previous speaker pointed out, we can combine mortality with health, not just look at the health of survivors and not just assume that everyone who's alive has the same quality of life. We can put those two together. And there are competing ways of doing that.

But any data set that has the SF-36 or

the 12-item subset can score quality-adjusted life years using available software. And I think that will be very useful in this field.

That metric, by the way, the red and blue that you just saw, has now been published in more than 20,000 peer-reviewed articles, including 2,000 well-controlled, randomized trials. But those tools actually had their roots in general population surveys. We developed those metrics to evaluate health policy. Will free care improve health? If we make people pay a large portion of the bill and we save money, will that harm health? So those surveys were actually population surveys, but people started asking to use them in clinical trials in the 1980s.

We got the idea, well, why do we have to change our concepts and our metrics for a population survey versus a clinical trial? And we began to push that envelope forward.

To use the metaphor from last month of a thermometer, those of us who cooked a turkey or a roast, we didn't use the same thermometer that we

use to measure the temperature of our body. We put

a cooking thermometer in the roast. It's a

thermometer that measures temperature at much

higher degrees. What we need to do in this field

is just to match the levels at which we're

measuring to where the population is scoring before

and after we change a risk behavior.

The advantage of using standardized

tools, as shown here, is that we can put smokers

and nonsmokers on the same metric that we put

chronic disease, that we put the well population.

And to carry this further, we can look within a

disease at different levels of severity.

So in validating measures for quality of

life assessment in tobacco research using the same

logic we did in the medical outcomes study,

published in JAMA in 1989, if we're going to look

at outcomes from people with a particular disease,

I want to know that I'm using a tool that can see

the difference between having the disease or not.

If not, how can I expect to see subtle differences

within the severity of the disease?

So as many of you have discussed today
and yesterday, we look at former, current, and
never-smokers, and we look at current smokers
differing in whether they're smoking a heavy or a
light amount.

Where they are on this -- and you can see
well and current and former smokers score pretty
high on the physical dimension, and you can see
over on the right, they score pretty high on the
mental dimension. And you can where other chronic
conditions that we're familiar with, the average
morbidity, we can see where that stacks up.

Given the interest of this industry, if
not the priority or the strategy of focusing on
otherwise well smokers, the ceiling effects that we
have with general population studies are even
greater. So one of the things we've been
addressing is whether we can raise the ceiling so
that people don't have a perfect score to start
with.

Now, the way we validate measures
is -- and what I'm going to say is in the spirit of

a vendor important point that I heard yesterday and

that I heard today, and that is, we don't want to

just get a result of quality of life differences

between two risk groups. We want to know how that

happened. What process caused that to happen? So

this is the kind of endpoint model that the FDA

likes when it's judging the results from a clinical

trial.

So we looked in the general U.S.

population using measures of the severity of

symptoms like I just showed you and using a

condition impact scale that's designed to be

condition-specific for quality of life. It's a

sample of all those content areas that you saw in

the previous table, but with an attribution to

smoking. Is your smoking limiting your social

activity?

The correlation between the severity

of the symptoms and the quality of life impact

attributed to smoking is substantial, .558. And

further, the worse the quality of life impact is

attributed to smoking, the worse the general

quality of life, particularly for the mental component.

So there seems to be a greater effect or association of smoking in the mental dimension, although it is clearly significant in both the physical and the mental.

We were able to replicate this recently in a MRTP trial, a small trial in Germany, using industry data. And this is just looking cross-sectionally at baseline. I'm staying completely away from any treatment comparisons here.

We see a positive correlation between the German translation of the symptom measure and the impact measure, and we see a substantial correlation -- negative, because quality of life impact, higher is worse -- and the general outcome is favorably scored.

Before I leave that endpoint model, I want to remind us that there's more to this continuum from a clinical, economic, and social point of view. These specific and generic measures are the best predictors of virtually everything we

want to predict in healthcare.

Here's a list: health in the future, how much it's going to cost next month or a year to treat you, job loss, return to work, work productivity, all the way down to mortality.

That excellent-to-poor item I showed you at the beginning, if you begin a five-year period with a rating of poor, you are 10 times more likely to be dead within three years than if you rated excellent. So even that rather crude, very subjective measure of the confidence or evaluation of your health is a very good predictor of mortality.

Now my last point is how psychometric methods are being applied to improve the rulers that we use to measure quality of life in obesity, in smoking, and in chronic conditions in general.

In 1983 in a health insurance experiment, we asked 25 questions about physical functioning. They were the best available at the time. We scored it on a zero to 100 scale, and the highest level of physical functioning that was measured was

whether or not you could climb a flight of stairs
without any limitation.

At the beginning of that trial, as we
published in the New England Journal, 75 percent of
the U.S. population that we later randomly assigned
to free care already had a perfect score. We
couldn't improve it. Well, maybe we could if we
had 300,000 people, but not with 7,000 people.
That's a ceiling effect. Fortunately, we had other
measures in that trial.

So when we fielded the medical outcomes
study, we reduced the number of questions for
physical functioning from 25 down to 10. But we
distributed them better throughout the range where
physical function varies in the population, and we
reduced the ceiling down to 30 percent.

We also abandoned the zero to 100 scoring
because if we keep lengthening the ruler and we
express everything as a deviation from the high to
the low, we keep distributing everything in the
middle.

So we did what has been done in

psychological testing for a hundred years. We

expressed everything as a deviation from the

middle, and we decided to have the middle be the

mean of the U.S. population. That has also been

done in 10 other countries. But we reduced the

ceiling effect with a tool, half the link, down to

30 percent.

In 2008, the first of the physical

functioning item bank studies within PROMIS, my

colleagues and I published an even better measure.

It takes only five items using an adaptive

methodology. The ceiling effect was reduced to

3 percent, and we have much more precision

throughout the range of the ruler.

There's still another paper, which I

think will come out this month or next month in the

Journal of Clinical Epidemiology, carrying that

even further because basically, all the physical

functioning measures have now been cross-calibrated

on a common metric.

So it's kind of like Fahrenheit,

centigrade, kelvin; whether you use a digital

thermometer or an analogue thermometer, like

Galileo tested, we can tell you temperature in

Fahrenheit. So now for the major domains of

health, we can tell you scores whether we use the

legacy tools or not. That's a huge advance.

Now, the model that makes that -- this is

my own model -- the model that makes that possible

is the item response model, shown in the little

diagram down on the right.

On the horizontal axis is physical

functioning, with a mean of 50; higher is better.

The vertical axis is the probability, not of

physical functioning; the dependent variable in

this model is which choice are you going to pick to

describe your physical functioning. And those are

the plots. That's the real probability curve for

each of the choices in the general population. And

those curves don't change in different populations.

Now, notice that the curves cross each

other, and that's the point at which we don't know

where you are from that item for sure. And that

threshold, which is a parameter of the item, is the

red bar you see on the right.

That red bar on the right is the red bar
over on the ruler. So these models give us the
marks on the ruler in the units of the quality of
life domain that we are measuring. They tell us
how to score the answer, and even more importantly,
they tell us whether or not to ask you a particular
question.

If you're nowhere near where we've
estimated you to be, we're not going to ask you a
hundred questions that we already know your answer
to. We're going to ask you the one that we don't
know your answer to from the model, and that will
reduce your confidence interval quicker than
anything else that we could do. So that's how we
can do a 40-item survey in five items in less than
a minute.

Finally -- this is my last point -- on
the ceiling effect. About 10 years ago, we started
studying disability, as measured in the neurology
literature. And with industry sponsorship, we put
together all of the headache-related disability

measures that were being used in drug evaluation or

in general in screening and monitoring patients.

          This is a representative sample of U.S.

adults who had a headache in the last month, about

a thousand people. We're administering a very

widely used five-item disability measure. The

criterion is a bank of all of the disability items,

all calibrated on a common metric. So we'll call

the latter the criterion.

          The correlation between this measure,

which has a lot of people with no disability, and

the criterion is .54. That's the product moment

correlation. And those people have no disability,

even though on the criterion they varied reliably

by three standard deviations, that's a ceiling

effect. Those same people were given a new

five-item measure that represented the full range

of the bank. In the same minute, that measure

correlated .94 with the criterion, didn't take any

longer, and it has no ceiling effect.

          So the solution to these ceiling

effects -- obviously we want to match our items to

the person -- is an adaptive approach to

measurement or a six-item HIT, Headache Impact Test

form, which is available on the internet at dozens

of sites or in doctors' offices on a pad. You

don't need adaptive testing. Six well-chosen items

can measure that range much better than what is

shown over on the left.

So let me say what I hope I've said very

quickly. I really advocate an endpoint model that

views outcomes as a continuum of the most

objective, hopefully closest to the cause of health

problems that we can measure in box number 1; the

human experience of specific symptoms that are the

life experience of that problem; the quality of

life decrements that are attributed to that

condition or to those symptoms; and the generic

endpoints over on the far right.

I think this field would do very well to

standardize the generic core. And those are the

eight most frequently measured domains that were in

the table that are represented in most instruments,

including PROMIS.

I would say that, to my knowledge, tobacco consumption affects at least seven if not all eight of those domains. It consistently affects both or is associated with lower scores in the physical and the mental component.

So standardize the generic core. Keep it short so that we can study other interesting variables, many of which were mentioned today. Raise the ceiling. Particularly with young, otherwise well smokers who get very high scores, at least at this time, make sure that we have room to measure improvements associated with change in their risk behavior.

That, of course, raises a question, well, how high up in health is really important? And that's an old issue that is very familiar to me. In 1986, there were anti-hypertensive medications that were safe, they were efficacious in terms of blood pressure control, and some of them had quality of life issues. And those were the more expensive ones. They didn't have quality of life issues. The free ones that they were replacing

did.

The issue was whether the quality of life differences that were observed in that trial -- I'm talking about Krug et al., New England Journal, 1986 -- that's the well-being at the ceiling that we're talking about right now.

It turned out that those were important. And if you have a 5 percent decrement in those for 30 years of treating hypertension, that is a huge loss of human capital.

So that's something analogous to what we're talking about now, I think, in terms of being able to measure the removal of a risk due to exposure that ultimately might shorten a life, but because most things that kill you, maim you first, it's going to show up first on the radar screen for quality of life. And we would be so much better off if we can know that answer long before the body count.

But we need to know what's an important difference, and we know how to do that. I have my opinions on that, but we'll talk about them some

other time.

Finally, we need to make data collection more practical. I favor self-administration. It's cheaper. It's more private. It's preferred 4 to 1. It's less biased. There are some people that can't self-administer.

Electronic data capture is a huge advantage, particularly on the internet. But don't program a 141-item static form. If you're going to use electronic data capture, take advantage of adaptive testing, which matches better the questions to the respondent.

You may answer different questions that me, but I can perfectly compare your scores. And anyone who doubts that can do an empirical test. Thank you very much.

(Applause.)

DR. WARE: Here are a few copies of the handouts. I don't have enough for everyone.

DR. DRESLER: Our next speaker -- and thank you very much for allowing the switch-around -- is Dr. Jia from Columbia University. He

is an associate professor there in biostatistics,
and he is going to present on to calculate quality-
adjusted life years from nationwide survey data.

Presentation – Haomiao Jia

DR. JIA: This is a study a little bit
different from most of your talks today. I heard
people say that we have a model. We do a lot of
tests. But we don't have data.

When I start the work on this
project -- actually it's almost 10 years ago -- I
talk to the people at CDC and they said, we collect
this nationwide data collection and have a lot of
data. But they cannot do the calculation of QALY,
and can you figure out a way to do that? So that's
the start of the problem. We have data and had a
lot of information, but the data was not designed
for the calculation of QALY.

So first of all, what is the quality
adjusted life years? It's a summary measure of
health that has included both mortality and
morbidity outcomes. For the mortality outcome,
this is measured by the year or life. For the

morbidity outcome, measured by performance-based quality of life. And that's the utility score Dr. Ware mentioned.

So performance-based quality of life score, it's a summary score that gives a value of one health status versus another, and have a zero for deaths and 1 for perfect health, and can be negative. That means worse, deaths.

So that means one year a person live in a score 0.5, that give you .5 for QALY, and same as if another person, and if only half a year but in perfect health. So you can add it up over the lifetime.

One way to do the calculation of QALY, it's you have RCT and you compare to treatment, standard treatment and the new treatment. And we can do this calculation of QALY from zero, time zero, to time R, that's end of the static. So the calculation, it's this part' X as the survival. Y is the quality of life score.

So that's QALY for a fixed period of time. But this calculation, sometimes can have

difficulty if they have a lot of people censored.

And the problem is the censor is related to the

health status. Because people have a disability,

they censor the data. So then you do the

calculation, maybe have a bias, have some problems.

Another why is more often used is

calculated life known QALY. So you calculate it

from a certain time to the end. Same problem if

the data, if you use cohort data, and if the data

have a larger number of censors, and the estimation

have bias, can be biased.

Also, this also have a problem. It's

when you do the comparison of two groups of people,

maybe smokers compared to nonsmokers, remember

this. They could be have a different age range.

So then you may have to consider and adjust for

age.

So from here, we say, may be easier to do

the comparison. It's for the QALE, the quality-

adjusted life expectancy. So then we give a fixed

age. So QALE means it's average lifetime QALY

start from a certain age, X, from the age X. So

you start calculate this.

So this gives you a good comparison. People at the same age, then you can compare the value. So you can compare smoker at age 18 to a nonsmoker at age 18. It can be both men. And you get an idea of how their life outcome looks like. So that's the quick summary of QALE and QALY.

The question is, why use QALY or QALE? And I said before, certainly is a measure of burden of disease. And the second reason is you can do it for some cost of illness and the cost-effective analysis.

Sometimes if you calculate a disease, and if this disease cause you loss, say one QALY, this disease of one QALY, then we can put a dollar value on that what it means is you lost one year of QALY. Right? So yes, we have the equivalent.

So this dollar value, yearly it's associated willingness to pay for the treatment of this disease. So that's for the economic analysis of this disease.

So next question is how to calculate QALY

or QALE. So when you do the -- for the
calculation, for this part, you need estimation of
quality of life score and survival function, or
maybe hazard function. And you probably want to
know the relationship between these two.

As I mentioned before, the difficulty for
this is the censored, the data. If you use cohort
data, most time they have a large number of people
censored at the end of study. That causes a lot of
problem. Particularly it could happen -- the
censor is related to the quality of life score.

So there is other approach. We begin to
think about this, and people propose. It's can we
estimate the score from nationwide, this cross-
section of survey? And then we figure out the
probability of death during each age interval. And
then we can figure out this probably easier.

So here is a summary of how do we
calculate. It's based on the life table. First,
we estimate age-specific, performance-based,
hazard-related quality of life score from
nationwide survey search as BRFSS. And then we

construct the U.S. life table from age-specific

death rate from the compressed mortality file.

That compressed mortality file we can get from

NCHS.   Then we combine these two.   We calculate

QALE.   So that's the idea.

So first step is estimate the

performance-based, hazard-related quality of life

score.   The problem is, most of the national data

don't have measurements, and the MAPS data have the

data for a few years, 2000 to 2003, have EuroQol,

EQ-5D questions.

Here is the start of our project.   The

CDC have quality of life four questions.   Include

one question general self-rated

health -- excellent, very good, good, fair,

poor -- and the three questions on the number of

physical unhealthy day, mental unhealthy day, and

activity limitation days in last 30 days.

These question have been included in the

BRFSS since '93, and NHANES since 2000.   So wanted

to know, is that the way use this?   It's four

questions, together utility score, utility value.

So we start thinking about how to convert from these four questions into the utility score. That convert to that have a big advantage. The one thing is then we have BRFSS data and can use that to do the calculation of QALY.

So there is a way to convert. We develop algorithm and relatively good validity, and the algorithm, it's not model-based. Okay? It's not model-based. It's nonparametrical. So basically, you base it on the answer for the four questions plus age, different groups. You get the estimate, the utility score here based on that.

Then the next step is to construct life table and calculate QALE. For this part, it's just the standard life table. I not use the complete life table; just to show you, it's an abridged life table.

So from compressed mortality file, we have number of people deaths and population in different age interval. So from here, we can calculate mortality rate by dividing deaths by population.

From here, we can make assumption of
constant probability of die during that interval.
So you figure out probability of deaths during that
interval. So you can convert. It's almost the
modified bias, and then solve for that age
interval, but a little bit different.

Then we start with hypothetical
population, say 100,000, and you multiply by the
probability die. During that interval, you get 653
people died during that interval. And then the
life years during that interval, it's this value.
That means that people not died, lived the full
length of that time. And for the people died
during that interval, they on average live about
the half, a little bit less than half the length.
It's also based on the assumption of constant
probability of die throughout the interval.

So we have the life years for that
interval, and all the way going down. Say it's the
last interval, the probability of death, is
100 percent because in the end, everybody going to
die. And you start -- in the last interval, you

have 40,000 people, and everybody's going to die. And the life year in the last interval also estimated based on the constant probability of die.

So then we just calculated the life year above that age. So you just add it up, go back. Then you divide it by the population, so you get life expectancy. So that's the standard life table, NCHS prepared.

So if we have the quality of life score estimate from nationwide survey for different age interval. And that's the value. We multiply this value to the life year, we get the QALY. Right? The QALY in that interval. And then we calculate QALY all the way down, and then the QALY above that age. So we just go back and then calculate QALE. This is the QALY divided by the population. So you get QALY at age 18, QALY at age 25, and go on.

So that's the quick summary of how do we calculate. And here involves the number of assumptions, and also we did some sensitivity analysis, try to look at violate that assumption, if that still hold. And it turn out it works fine.

Say, for example, the quality of life score is heavily related to the probability of die. It still works fine, and so on.

So the next thing is the QALE loss due to smoking or due to any disease. So in the literature, the definition is a little bit unclear what it means, the QALE loss. And basically, it's two kinds of loss.

One, I call the individual QALE loss. That means if we compare the smokers to nonsmokers, the smokers have a small QALE compared to the nonsmokers. So they have a loss. So that loss, we call it the individual QALE loss. So basically, if we calculate QALE by smoking status for smoker and nonsmoker, the difference -- so nonsmoker have a bigger QALE, right, and the smoker have a smaller QALE. So the difference is individual loss.

Another type is also popular to use. We call it the population QALE loss. Basically, it's the difference between the QALE for entire population and QALE for nonsmokers. What's this mean?

This means if all the smokers actually
have the same mortality and morbidity value, the
population, entire population, will gain QALE. So
that's a little bit more about how the impact on
the prevalence of smoking.

So then the problem is calculate QALE,
but how to calculate QALE by smoking status. It's
an easy one. You can calculate the quality of life
score by smoking status, very easy, but the death
rate by smoking status is not available in either
estimate.

So here you need a little bit of
modeling. We have a hazard ratio, estimate a
hazard ratio, and the prevalence of smoking. And
the death rate, that's relatively easy to get.

Also, we also did a sensitivity analysis.
Say if we get a poor estimation of prevalence, does
that change the difference? Also, if we get a poor
estimation of a hazard ratio, can we get also a
similar result? And it turned out actually it's
pretty stable. If we get very bad prevalence
estimation by, say, 100 percent, still you get a

relatively stable result.

Here it's the hazard ratio. Here is some result. The last part is the QALE for smoking. For smokers and nonsmokers, certainly the loss [indiscernible] throughout the different age group. And so you can do the life expectancy difference.

This is the difference between men and women, and you can see men number is a little bit higher. But it's about the same, similar pattern.

So here we look at the trend. That's the advantage of BRFSS data. You can look at it from '93. So the smoker and nonsmoker QALE changes over the time, and the QALE loss.

QALE loss actually increased. Okay? Over the time, it increased. So the question is why it increased. We think, well, we have got better. Right? But remember, the QALE loss is nothing to do with the prevalence of the smoking. It's related to -- compared to smoker and nonsmokers, the difference just become wide.

Then we examine carefully why they

increase. We look at which part contribute to the increase. Actually, the increase, it turned out it mostly is coming from -- it's a loss due to mobility, not due to from mortality. So mortality actually part. It's not really change. The change part, it's the quality of life score change.

So that's actually sounds make sense. Still, it's a difference. The year of life lost not really change much, but your quality of life change. It could be because people quit smoking, become former smoker. Here are our definition of smoking. It's current smokers.

So let's just skip the next few slides. And we also can calculate it by state. That's the advantage. And we look at a number of bias, and use different data, MAPS data and BRFSS data. What's the difference? It looks like -- data set.

So for this study, I think the big advantage is we can use -- this is called the legacy data, the BRFSS and NHANES. Then we can go back 20 years before, all the way to 1990, to look at a trend.

The reason to do this is the CDC realizes they are not going to be able to put utility questions into their large survey. So they have to rely on these older questions. That's the main reason to do that. So that gives us a good chance to track the change over the time, and also look at the difference between the states and even local area.

Actually, we did a similar analysis for New York City in the last -- since '93 also because New York City have this big campaign of smoke and preventing smoke. So we can also get that. So that's the purpose for that.

DR. DRESLER: Excellent. Thank you.

(Applause.)

Discussion - Panelists

DR. DRESLER: Could we have the speakers come up, please, and Dr. Erik-Rutqvist.

Dr. Erik-Rutqvist from Swedish Match, do you want to start off with some comments, please?

DR. RUTQVIST: Yesterday, Mitch Zeller, in his introductory remarks, referred to the public

health standard that CTP uses to make decisions.

But as we all know, there isn't much in the statute

or in any guidance documents about what this

standard is.

So as an applicant, what should be your

conclusion from that? Well, my take is that you

have to be innovative and come up with something

that makes sense to you, and hopefully, then,

convince CTP that it's a good idea.

Now, Swedish Match intends to submit a

modified risk application within the near future.

And the central product-specific evidence in that

application will obviously be the epidemiological

data based on the Swedish experience, clinical

trials, premarket research, and also results based

on modeling, which we believe is valuable both in

the premarket and postmarket setting. And the

model that we will use is the dynamic population

model that was presented by Annette Banchand

previously today.

Now, in its current form, the dynamic

population model looks at overall survival. But

intuitively, it makes perfect sense to include also some sort of measure of quality of life, as has been illustrated during this session.

That said, I think we should always remember that survival will always remain the basic measure because there can be no happiness, there can be no quality of life, if you are dead.

Now, including some kind of quality of life measure is one of the possibilities that we are considering incorporating into the dynamic population model. But there are some conceptual and even ethical issues with measuring quality of life. Some of those have been touched upon today.

For instance, which should be the domains that we measure? And it's obvious that if you use the WHO definition of quality of life, domains related to mental and social functions should be given more priority than in some of the instruments that are used today for measuring quality of life.

Also, these domains may change over life. How do we deal with that problem? And who should assign utility scores to these different domains?

It's also an ethical issue, I think. For
instance, using a one-dimensional metric to
summarize such a multidimensional concept as
quality of life, who should assign the utility
scores? Should it be people who have experienced
the conditions that we try to measure, or should it
be the general public?

Also, some people may have difficulties
with a metric that equates a rather trivial,
perhaps, condition in a large part of the
population with severe outcome for a small number
of people, even death. I think such a concept,
there could be problems from an ethical point of
view with such a concept.

Then again, intuitively it is reasonable
to incorporate some form of measurement of quality
of life because it's quite clear that just
measuring survival probably doesn't capture, or
definitely doesn't capture, all the morbidity
associated with smoking. I think I'll stop there.

DR. DRESLER: Panelists, did you want to
respond back to this? Any of the panelists? No?

As we get ready to open up for questions,
I have a couple because this is one that I was
thinking of earlier. And Dr. Jia, you had started
to address it, and I thought the other two
speakers, too, and that is, looking at how do you
assess the cost to it? How do you put the dollars
to that?

So DALYs, and particularly what I've
looked as QALYs. So you had some of the cost
analysis, Dr. Jia. But I'm wondering if you can
look at how do you use DALYs? You could use QALYs.
You could use dollars. And if you could address
that, please.

DR. JIA: The reason I -- well, I didn't
start with today's topic. Actually, we start with
the talk with economists at the CDC. And the CDC
have developed these four questions since '93 and
put it in the BRFSS.

Every time they push the people to use
it, to calculate some kind of cost-effective
analysis or cost of illness, and the economists,
their problem, it's, well, we cannot use it. You

have to some kind of convert it to a certain way to the utility score. Then we can make analysis of the impact.

Say, for example, this is one thing they come up with. Say CDC, we give out money to some state to do some intervention. And they wanted to know, yes, it looks -- get a good result, is a smoking cessation program, and the prevalence decline, and mortality decline, and related to smoking. And is that worth it? Because it costs a lot of money to do that.

So that's the conversation beginning. We don't have unlimited resource, the money, where to spend. So that's the reason to do this type of calculation of QALY or QALE.

Regarding to how much the dollar and the willingness, it's typical that people choose $50,000 per QALE gain. And it doesn't have to be that number. It can be other numbers, some little bit high, some little bit lower. But that's not the point. The point is once you do this type of analysis, then you can compare to different

program, which one is most cost-effective.

DR. DRESLER: Yes?

DR. WARE: Both speakers made a very good point about which domains and the issue of content. And also, you made the very good point that it was kind of a historical accident that the utility tools that are used in economic evaluation have a different descriptive system for the states of health than the psychometric tools that are used in clinical studies and so on.

It makes no sense to me that the domains of health that are important to the public that are the burden of the disease and the benefits of treatment should be a different set of domains if you're going to divide it into dollars than if you're going to do something else with it.

So the EQ-5D, for example, the 5D is five domains. It's really actually four because two of the domains are two levels of the same domain, physical functioning.

So it leaves out four or six of the eight or ten most important domains that patients value

and that are just consistently part of the

descriptive system for most therapeutic areas. And

the FDA has been emphasizing in its guidance

document the importance of content validity.

So we ought to have a common descriptive

system, and we should have a way of evaluating its

comprehensiveness. I don't understand why some

things would be left out in a utility-based

descriptive system.

Of course, that's why I like the SS-6D,

because it has more content in it. But there are

things missing from that. It doesn't have sleep

adequacy. It doesn't have sexual functioning. It

doesn't have cognitive functioning.

So we're leaving a lot out as it is. But

to leave out four or five of the ones that are

consistently, in virtually every clinical trial of

PROs, affected by the disease and are a benefit of

the treatment, that's an incomplete economic

evaluation.

If you happen to have one of the diseases

that most affects the left-out domain, as was

noted, often emotional and social and real
functioning, that's a bias that just isn't fair
from a societal point of view.

DR. NG: So in terms of how to attach
costs, link costs with DALY, there are two ways.
One is you can do econometric analysis, so you have
a population intervention. So you can calculate
the changes in DALY per dollar. In one of our
health financing reports, we do look into those.

Now, the other way to think about how
costs can be incorporated in the calculation of
DALY is, remember, there's the counterfactual
distribution. In comparative risk assessment,
there are different ways you can assume the
counterfactual.

One is to think about what is the
possible, most cost-effective measure to control
tobacco, for example. So based on the cost-
effectiveness strategy, you can have that as your
counterfactual and compare. If that counterfactual
is true, if we have such policy, what might be the
attributable burden of smoking? So there are two

ways that we can conceptualize costs with DALY.

DR. PHILLIPS: So -- sorry, Carl

Phillips, CASAA. So we had a theme here of trying

to more precisely estimate the damage being done by

smoking. And I'm as interested in just finding

things out as the next person. But the description

of the workshop is for regulatory guidance and so

forth. So I'm just wondering, is there anywhere

within improving these predictions, these measures,

that you actually see a difference in smoking

policy?

Because my observation is that when

somebody identifies a potential policy that they

think will reduce smoking, the next phrase is

always, "And therefore we should do it," or, "If we

can get away with it." The next phrase is never,

"It's a really close call between how much cost

that imposes on society and the benefits it would

bring about, so we need a better estimate of how

much damage smoking is doing before we decide to do

that."

So I'm just putting it to you. What

could be done differently if we got better

estimates along the lines of what you're thinking?

Then to halfway jump ahead and propose

an answer to see if you'll buy into it -- and this

is mostly to Dr. Ware -- once you eliminate most of

the costs from tobacco use by not making it

smoking, then the benefits start to stack up pretty

strongly against the costs with a low-risk

alternative.

Do you see your research being able to

tease out the net benefits of using a low-risk

product as compared to abstinence?

DR. WARE: Well, aren't the quality of

life decrements with lighter or cessation of

smoking the benefits of an alternative to that? So

we want to be measuring those because that's the

gain with the alternative product.

The other thing is, in reading the

smoking literature, a number of articles I've read,

there may be a role in appealing to the public of

helping to better understand not what's going to

happen in 60 years, which many people discount to

zero, but the decrements you have right now in the

quality of your life might be part of the appeal.

Just like in dealing with adolescents

long ago, their hair quality and their complexion

and other things were the major things that

motivated change, not survival. They have no

concept of what's going to eventually get them.

So I think there may be some value in

helping the public to better understand what life

is like with smoking, and not just the implications

for longevity and the later morbidity with the

chronic conditions.

If I can make one more point -- because I

may not get to make another point -- everything

that we have seen as far as the differences between

current and former smokers are two to three times

larger among the chronically ill, controlling for

everything else in the model, particularly in the

emotional domain.

So evaluating our products only in

otherwise well smokers, we're not going to

understand that there's a huge gain to be gained in

life among people who already have a lot of

disease-related morbidity, some of which is

smoking-related and some is not.

So I think that needs to be replicated.

But it's a pretty strong interaction when you

compare current and former smokers who are well

versus chronically ill. The differences to the

advantage of former smokers are about three times

larger for those with one or more chronic

conditions, using standard epidemiology.

DR. DRESLER: We'll just take two more

questions because I just find this really

interesting. I don't want to quite quit. In the

back.

MR. SAXENA: Hi, I'm Kunal, a graduate

student from Virginia Commonwealth University. I

have two questions, if I may. My first question is

for Dr. Marie.

Going back to the slide where you

compared DALYs with quality-adjusted life years,

don't you think these conditions specific to

smoking, more precisely conditions like lung cancer

or COPD, et cetera, are more inclined towards the

quality of life rather than the disability of life?

DR. NG: Disability of life in terms of

the weight is actually -- it's related to quality

of life. In fact, the way that disability is

calculated, it does have that component of quality

of life, except that we are not measuring that

general health status.

So this is one of the key differences

between QALY and DALY. QALY tends to be more

general, as an overall feeling thing. But DALY is

very disease-specific, specific to that condition.

So this is one of the differences. But they can be

very related. So it's not a complete distinction.

In fact, you can also see QALY that's very disease-

specific. And so it is a very confusing

distinction in some ways, I have to say.

MR. SAXENA: Thank you.

My second question is for Dr. Ware. When

you're using the instrument -- for example, an

instrument to measure the quality of life

scores -- do you use the same technique that you

use to measure the quality of life score for a

person suffering from a smoking-specific condition,

and then use the same method to estimate the scores

in a person suffering from the same condition, but

the cause may be something different? Does that

make sense?

DR. WARE: Yes, it makes sense. A very

good question. For the general part, that's what

allows us to compare the burden of different

diseases and the benefits of different treatments.

There's actually an article in review

right now reviewing the 17-year history of PROs in

randomized, controlled trials, a rank order of the

effect sizes by 14 therapeutic areas. That's what

you can do with a general measure. That's box 4 in

that diagram you saw.

Box 3 is the other one that you

mentioned. And we need to spend some time, and we

have good methods for determining the ability of

patients to make valid attributions. These are

patients with multiple chronic conditions. Do I

know how much of my pain is my arthritis versus my

coronary disease, and so on.

As the symptoms that define these start getting the same, like breathlessness -- that's a symptom in COPD, asthma, CKD, and other conditions. So it's not really specific. There are no specific symptoms that are so specific that if they change, we'll know what caused it. So that's why we do randomized trials.

So we need to push the envelope on the ability of the public to attribute, in the face of multiple conditions, the impact on their quality of life. And actually, I think we're going to really be surprised at how well they can do that, but there are certain clusters or syndromes that it's just kind of like mush, and they really can't sort it out any more.

But still, knowing that assessment, even though they don't know the ideology as well as three specialists would, that is a very valuable predictor of everything you want to predict. They just don't quite understand the causes. But bad is bad, and everything associated with that is bad.

DR. DRESLER: We have one more question.
Yes, up front, in the second row.

MR. PERKINS: Roger Perkins, National
Center for Toxicological Research, FDA. My wife's
a psychologist, and she would scold me if I didn't
ask this question. I think I will address it to
Dr. Ware.

I didn't hear much, in talking about
measures of quality of life and even endpoints,
about socioeconomic status. Because it seems to me
when we talk about quality of life, it's hard to
divorce that from some kind of stratification in
our models and so forth; also, in trying to design
policy.

For example, if you tax tobacco, which I
think most states do and pretty heavily, it has
some mitigating effect on tobacco use. But it
places a real burden on those at the lower strata,
and who knows, that first cigarette to that last
cigarette of the day may be the high point of their
day.

So how do we factor socioeconomic status

into measures of happiness, quality of life?

DR. WARE: Well, there are several books from Michigan during the peak of the social indicators movement, probably in the 1970s or thereabouts. Most of them have an apple on the cover for some reason, quality of well-being, whatever.

But that's that list of eight or 12 things that I said, one of which is health. The other has to do with safety in the neighborhood, and transportation, and economic considerations. So those are very important definitions, if not determinants, of quality of life. And of course, there's a lot of income/health relationship.

But the first work that we did with quality of life measurement at Rand for the health insurance experiment is we put health in the forecasting model for expenditures. And it was more than an R square thing. It changed the coefficients for the other variables in the model, including income.

In other words, we were attributing to

income something that was equally attributable to
health. And those two probably go in both
directions, downward mobility due to health
problems, and income limiting your resources and
getting the care that you need.

The other thing is, it just underscores
the importance of us having good representative
samples when we do this work, including clinical
trials of new modified risk tobacco products. The
speakers said in modeling, exporting or migrating
the predictions to the population at large and not
just some little corner of the world that we
surveyed or analyzed.

But congratulations to your wife for
bringing up income.

DR. DRESLER: All right. Thank you very
much for this very interesting panel and the
discussion.

Yes, we are still behind. But we don't
think the last session is going to go as long as is
on the agenda. So can we please take a 10-minute
break and then come back, and let's talk about what

future directions will be.

(Whereupon, a recess was taken.)

DR. DRESLER: George Rochester, again he's our branch chief for statistics and future directions.

Moderator – George Rochester

DR. ROCHESTER: Future directions. Chief of future directions. New titles.

All right. So we are on the last leg of our relay. So we're getting done, and it's nice to know our stadium is not empty. We still have people here, so this is wonderful. Great participation so far.

I think at this time I would like to invite our three panelists up. I'll ask that each one will introduce himself and then say a few opening remarks, if you wish, and then we'll have a robust discussion following that.

So Conrad, you want to go first?

Discussion - Panelists

DR. CHOINIERE: Good afternoon. My name is Conrad Choiniere. I am the director of the

Division of Population Health Sciences here at CTP

within the Office of Science. I oversee the

statisticians, social scientists, and

epidemiologists.

I had a few thoughts to share after

hearing the very interesting discussions over the

past few days. Some of them are future directions.

Some of them may seem more like flights of fancy.

But before I go too deeply into that and

get into trouble, I should probably remind everyone

here that the views that I'm going to express now

are solely my own, not FDA's, and they do not

represent any official agency position or policy.

So yesterday, and today, actually, there

was a lot of talk about how the Act lays out a

public health standard for tobacco products. Now,

the Act is pretty vague in its description of the

public health standard. It tells FDA what it

should consider in order to determine whether or

not a product meets that standard, but it doesn't

provide many specifics about where the bar should

be set for meeting that standard. And I'm sorry to

disappoint you, but I'm not going to provide you

those specifics today.

(Laughter.)

DR. CHOINIERE: But what I do want to

talk about a little bit is what we're pretty sure

that standard is not. I should have explained a

little bit that I've spent much of my time here at

FDA assessing the impacts of risk communication and

marketing on consumer perceptions, beliefs, and

behaviors around FDA-regulated tobacco products.

One of the things I've worked on heavily here at

CTP is the modified risk tobacco program. So many

of my comments will have that as my perspective.

So when talking about the public health

standard in the context of modified risk, it's

clear that it's not merely about reducing exposure

or the risk of harm in an individual user of a

tobacco product. We really need to consider the

impacts of a modified risk tobacco product on the

health of the population as a whole.

So yesterday, there was a lot of

discussion about what the scope of modeling should

be for modified risk. And there was a good answer.

Models should address morbidity and mortality

related to tobacco use. But in order to do that,

the models need to address tobacco use behavior,

which includes initiation and cessation.

But how do we use models to predict

initiation and cessation? How do we use models to

predict who is likely to use a product that has

never been marketed before?

Now, I am trained as an economist, and as

an economist, we spend most of our time building

models. And from my experiences with models and

the ones that I've built and used, you can build

very good models that describe what has happened in

the past, and you can use them to test hypotheses

about why things may have happened the way they

did. Or you can use them to compare policies. But

they aren't always particularly good at predicting

the future.

Could we have had a model to predict the

popularity of e-cigarettes, for example? And how

can we use what we have learned in the past to

actually predict the likelihood of product
initiation in the future?

So here are a few thoughts I had about
the directions we need to move in, in order to get
to that point. The first area I thought was -- and
this has been said many times over the past few
days -- and that is to improve data collection.

We heard yesterday, and it's said often,
garbage in/garbage out. So clearly we need to
improve some of our data sources and actually the
types of data that we're collecting.

We have some studies that will be going
out soon, if they're not already out, such as PATH
and others, that will provide us a wealth of
information that we can use in these models,
particularly in the areas of tobacco use
trajectories and transitions.

But we don't only want to know what the
transitions are, or the likelihood of transition.
We really need to know the mechanisms behind those
transitions to understand why those transitions
exist to begin with.

So in order to do that, I think we really need to do some more work collecting information about the "agent." And I'm putting it in quotes because we've used the word agent in different ways over a few days. In this, where I'm talking about the agent here, I'm using it in the economic sense, the consumer of the products, the decision maker.

What are the underlying mechanisms from the consumer perspective related to product adoption or use or experimentation? Can we better identify the characteristics of the individual that make them most likely to experiment or adopt or use tobacco products or even types of tobacco products?

There has been some work looking at characteristics such as risk affinity or social networks, but there are other factors that we potentially need to be looking at and collecting information on.

Once we understand that, can we then identify how many individuals exist in the population with these characteristics so that we can better predict the impact of product initiation

into the market?

There's also talk -- and I think it was yesterday -- Gary Giovino mentioned incorporating the "agent." Again I'm putting this in quotes because he was using agent in the epidemiological sense, which is the tobacco product itself. Can we incorporate information about that agent into the models?

Here I'm going to get into my flight of fancy. So economists have these models called hedonic models. I think some of the economists in the room have heard of these. And these are models that can be used to estimate a consumer's willingness to pay for a product based on the features of the percent.

So for instance, we know a lot about real estate. We know a lot about what people pay for houses. And we can put in the models all the various features of houses, such as the number of bathrooms or the number of bedrooms, how large the house is, if there's a dump nearby, and other features of that. And we can predict with pretty

good accuracy how much someone is willing to pay
for that house.

So maybe -- and I'm not saying that we do
this for tobacco products -- but maybe there's some
analogous type of thing that we can do with tobacco
products.

Wouldn't it be nice if we had something
similar for tobacco products where, if we knew what
the various features of tobacco products were that
drove consumers to perhaps experiment with a
tobacco product, then perhaps we might be able to
do a better job of predicting whether a new product
would be likely to be adopted, or if a new
product -- which having a modified risk tobacco
product adopted by a smoker may be a good thing.
But having one adopted by a naive non-tobacco user
may not be a good thing, so understanding the
interaction between the type of individual and the
features of the tobacco product to get a better
understanding of how likely it is to be used.

Which brings me to my last couple of
points, which are related to getting more about

understanding the reasons for using tobacco

products, and also the issue of transparency that

was discussed yesterday.

Now, I notice that on many of the

models -- with possibly the exception of the Sandia

model, and maybe I was missing something -- many of

the models had very simplified characterizations of

consumers. And I know we've talked about parsimony

and making models simple. But there's more to the

consumer than their tobacco use behavior and their

demographics. And those aren't necessarily the

driving features or characteristics of individuals

that lead to tobacco use.

Marketers typically stratify individuals

in different ways, perhaps based on the

psychographics of an individual. And these may or

may not be related to demographics. But perhaps we

might want to look at ways to incorporate

psychographic elements into models.

The opinion dynamic model presented by

Sandia may have attempted a bit to do that, but I

think more work needs to be done to more

effectively incorporate that type of information into these models to predict outcomes from tobacco product marketing.

Which gets me to some transparency, and that is that in this room we have representatives from some of the most profitable companies in the world. And I would bet that these companies rely on some pretty sophisticated marketing techniques to understand the consumers and forecast the likely success of these new products.

So perhaps we can learn something from these companies, and perhaps they have data that they can share with the tobacco research community to provide insights in this area, and that we can then develop some marketing models or adapt marketing models to better understand the public health impacts of tobacco products. Thanks.

DR. MEZA: Hi, everybody. I'm Rafael Meza. I'm assistant professor from the Department of Epidemiology at the University of Michigan. I am a modeler. I am a member of the CISNET lung group; in fact, I just became the coordinating PI

of the CISNET lung group. So I have a lot of

experience on modeling smoking, smoking behavior,

and particularly its relationship to lung cancer.

Maybe just a few remarks from a modeler's

eye in terms of what I've seen and heard today.

And I apologize because I was not able to be here

yesterday. So although I missed a lot of the very

good talks that happened yesterday, through

conversations with all of you today and just the

remarks already presented, I get a sense of many of

the things that were said.

Some of the things that you've probably

realized is that we have many types of models,

ranging from dynamical systems to individual-based

microsimulation to agent-based models. And the

point that I wanted to make is that that's a good

thing. It's not bad to have many models.

So in some sense, the more models, the

better, maybe with some caveat to that, because as

was said by Dr. Choiniere --

DR. CHOINIERE: Choiniere.

DR. MEZA: Choiniere; I apologize -- by

Dr. Choiniere is that different models will get

different points or are going to be good at seeing

many of the different important issues that are

related to new tobacco products and developing

policies to regulate and control.

Some models will be good to capture

population, general trends of behavior. Some

models will be really good at capturing the effect

of behavior and how this behavior leads to the

adoption of a new product, switching, or then

cessation of smoking and maybe stop using of any

other product.

So different models that have different

scales will be good at capturing different patterns

that will be relevant to the real questions. So

that's the main thing.

That being said, it's also true that

sometimes you want to have maybe a multi-scale

model that's trying to connection all the different

issues from individual decisions to collective

behavior, social networks, eventually leading to

population patterns. It's a good thing also to

have. So there is that.

Other things that I wanted -- and this is
a reiteration of the points that have already been
said -- is that as we've seen, there are many
models that address patterns of tobacco product use
by maybe gender, maybe by age, and maybe other
factors. Socioeconomic status is something that
many models don't have, and it may be an important
factor to include in many of these models, as well
as race.

Part of the reason -- and parsimony is
one of them -- we want to have the simplest models.
But it's also, as it was already mentioned, the
lack of data. And that's where many of the efforts
that I think should be made are to, actually, the
collection of data that allow for the prediction
of better models.

One thing that I think is important is
that good questions lead or facilitate the
development of better models. So in some sense,
having a guidance of what are the relevant
questions that need to be addressed is a good way

to helping modelers who may want to develop tools
to help address these questions do a much better
job. So in some sense, clarity and a clear
direction of what are the issues that should and
want to be addressed is something that is very
helpful.

I completely agree that participation of
the industry is very important because, one, for
the side of information, having access to
additional information that probably is not
available just from cross-sectional surveys or
prospective cohorts that may or may not be
accessible.

So bringing them into the picture is, as
well, important because they are going to be able
to provide a lot of information that can be used
for -- again, I'm repeating myself -- developing
better models to address the important questions
that we want to do.

I also wanted to say, going back thinking
about the presentations today -- and particularly,
I really liked the points made by David Levy this

morning in terms of the use of models to address

different aspects of policies. Right?

So using models to address or try to

assess the efficacy of past policies -- what's been

the contribution of taxes versus clean air laws

versus other different policies that -- what's been

their relative impact into the trends, in this case

decreasing trends, of smoking that we've seen.

We're also using models for forecasting,

not necessarily for forecasting what's going to

happen, right, like use models as a way to try to

determine exactly what's going to happen in the

future, but just trying to use them to try to get a

sense of what if we impose this policy? What if

people start behaving one way versus another?

Which is not necessarily what's going to happen,

but as it was said also in another talk, it's just

to get us thinking about the mechanisms and the

procedures that may lead to some behavior that we

actually want to have.

The third impact or the third use of

policies -- or, sorry, of modeling -- in terms of

trying address policies is what David called

heuristics. Right? Or again, trying to understand

mechanisms or theories about them.

Again, I think that was a very good

summary of the way in which modeling could help to

address the development of policies for the

implementation of these new tobacco products. I'll

stop there.

DR. HAMMOND: Great. I'm Ross Hammond.

I'm a senior economist at the Brookings

Institution, where I also direct a research center

that does a lot of complex systems modeling. I

have a lot of experience in doing policy-relevant

modeling, although mostly from outside the tobacco

field, which I'm relatively new to.

I want to offer three thoughts. The

first, which you heard me say yesterday and which

all my colleagues have also repeated, but I think

repetition is a good thing in this instance, is to

highlight again the importance of a diversity of

models. And I want to slice this a little bit

differently by thinking about three different

dimensions of what it means to have a diversity of models.

The first is to have different methods of modeling that are addressing the topic space. And within computational modeling, that would include both top-down and bottom-up models. It would include non-computational models, analytic models. It would include statistical models. And we've heard some of that diversity throughout the two days, and I think that's important to encourage.

The second kind of diversity is the diversity of different parts of the system or process that one might model. And I think Doug Luke made a very important point earlier today that we shouldn't require any and all models to have the same endpoints or to look at the same pieces of the process. And sometimes different models have different strengths, and we should embrace that and allow exploration of many pieces of this complex space.

The third kind of diversity is even among models that are using the same method, that are

looking at exactly the same topic, that are using
the same data, there can be real value in
independently developed models, multiple models
developed independently that are looking at the
same topic, even with the same method.

We heard some discussion of that earlier.
CISNET is a great example of that. I've been part
of infectious disease modeling, that work called
MIDAS, that also has that same strategy. And I
just want to underline that where policy-relevant
modeling is going on, that's a really important
source of diversity.

So it's not enough to have methodological
diversity or diversity of topics, but it's also
important to have multiple independently-developed
models that are addressing the same topic with the
same method.

The second thing that I want to highlight
is that I think there's a real role for both
funders and the consumers of policy models, the
policy-makers, the regulators, to really help
facilitate this kind of diversity, both by putting

forward a diverse set of questions and making

available and facilitating access to the kind of

data that these kind of models require to allow

common sources of data to be used across models,

and to explicitly encourage comparison and

crosstalk between models and modelers. So that's a

second point.

The final point is that the frontiers of

this kind of work that occur to me, besides this

application of new methodologies in this space that

haven't been as widely used, is that there's an

exciting explosion these days of new data sources,

whether it's EMA, which was mentioned earlier,

social media, which mentioned earlier.

There's a huge discussion going on in

other parts of the regulatory space about big data

or broad data. And I think the intersection of

these new kinds of data and these new kinds of

modeling tools is a very exciting place for this

group to consider.

The last thing that I'll say is that I

think a lot can be learned by looking at how

modeling and modeling that's policy-relevant has

been approached elsewhere in public health and

outside of public health.

In particular, some strategies that have

been used in other parts of public health to bridge

this above-the-skin and below-the-skin barrier and

to get connection across those different system

levels might be very relevant for the tobacco

context. So I'll stop there.

DR. ROCHESTER: All right. So first,

maybe, is there anything that's burning from what

you've just heard from our three speakers that you

would like to perhaps -- go ahead.

DR. BEACHER: Hello. I'm Felix Beacher.

I'm from Philip Morris, and I have a question for

Conrad Choiniere.

That is, you talked about predicting

usage behavior from premarket data. I think, in

Philip Morris, there's an uncertainty about exactly

what that means, what kind of analyses need to be

conducted, and also whether it's actually a

requirement as part of the MRTPA. Could you talk a

little about that?

DR. ROCHESTER: Did you hear the question?

DR. CHOINIERE: No. I heard the question. I actually think that -- I'm not here today to provide guidance on how to complete a modified risk tobacco product application. So I don't know if it's appropriate for me to answer that question.

All right. Then the first, which would --

DR. ROCHESTER: Could you remind us of the second?

DR. BEACHER: My concern is simply the vagueness about what -- what information we've had here is potentially a block to us conducting any analyses at all. And the second question that I had is whether or not it's a requirement or what the advantage is from the point of view of the tobacco company.

DR. CHOINIERE: Well, modeling in and of itself is not a requirement. It's not in the

statute. It's a tool that FDA has suggested that
could be helpful for -- we put it out there first
in the modified risk tobacco product guidance. But
I'm sure it would be useful in a number of areas in
tobacco control. Its venue is in tobacco control,
and we thought it was applicable in the realm of
modified risk.

So yes, it is a bit vague. I mean, we
are talking about looking to the future. And in
order to determine if a product is going to be
appropriate for the promotion or the protection of
public health, we need to try and make some
assessments of what the marketing of that product
will have on the population as a whole in the
future. So it seemed to us that modeling may be an
effective means of trying to get some of that
information.

DR. ROCHESTER: Another question? Go
ahead.

DR. SOLYST: Jim Solyst with Swedish
Match. And staying with Dr. Choiniere, Conrad, you
stated that here in the room there was

representatives from profitable companies with

marketing expertise data and models.

Before I heard that, I was prepared to

summarize what I feel is industry's product

stewardship obligation in the modeling and

statistical arena. And certainly that obligation

should be to contribute to the scientific knowledge

of models, and to be willing to, as my colleague,

Dr. Rutqvist said, be innovative. Be innovative in

model development. Be innovative in the definition

of public health benefit, not wait for CTP to tell

industry what to do, but take the steps necessary

that industry feels.

As part of that, certainly, be publishing

the scientific literature, and find other ways for

peer review and transparency, such as presenting at

these type of meetings, and SRNT, and hopefully

others.

But back to your excellent suggestion or

comment about just simply using the marketing

expertise. I think in our experience at Swedish

Match, we've had to work with our marketing

researchers as we developed premarket consumer

perception research, for example. They do have

great ideas. This is their job. They don't know

the regulatory context, and hopefully they're open

to that. Sometimes they are, sometimes they

aren't. But it's a great idea.

I think everybody wants the relationship

or the environment for discussion to move forward,

and this would be a perfect example of industry

being willing to share their knowledge and models

on marketing that could address, beyond just the

regulatory context, some of the issues that you

brought up during your presentation.

I don't expect CTP to say anything more

on that. I think it's up to industry, to

individual companies, to propose ways that they

might be able to share that type of information

outside of even an MRTP application.

DR. CHOINIERE: I just wanted to make a

comment on that. I wasn't suggesting simply using

marketing expertise, but I was suggesting as a

possible way of expanding the way we look at how

we market -- not market, but how we want to model

tobacco use or the likelihood of adoption of a

tobacco product, maybe just another perspective

that we might want to incorporate among all the

various perspectives that we already have in the

modeling.

DR. ROCHESTER: All right. Well, that

brings us perhaps to -- you've got a question?

DR. LEVY: I have a question.

DR. ROCHESTER: Great.

DR. LEVY: And I think this might be the

most important question that we're going to have to

address and what's going to make the most

difference in the results we get from modeling.

And that's a simple question. What is use? What

is use?

Oh, I'm sorry, David Levy. I'm from

Georgetown University.

So the question is, what is use? When we

talk about use, we think about it in terms

of -- usually you're asked how many times did you

use this product in the last week? How many days

did you use it in the last week? Or so on. These

are important questions.

I'd be willing to bet that I could find a

definition of e-cigarette use, which would show

that e-cigarettes is a gateway, and find a

different definition, which would find that it's

not a gateway. That's how critical I think this

question is.

It becomes an even more difficult

question when we talk about dual use. What does it

mean -- you're a cigarette user? What does it mean

that you're also using e-cigarettes? Does it mean

you use it every day? Does it mean you use it once

a day? What does it mean? These, I think, are

going to really be the most critical questions that

are going to determine the impact.

Before, there was a mention of the

importance of marketing. I think we're going to

have to look at the marketing literature. I think

that's the literature where they've really tried to

look at these issues. And these are going to be,

in my opinion, the most important issues that are

going to determine the outcomes of our modeling
efforts.

DR. ROCHESTER: Yesterday, for instance,
someone brought up the issue of having some
consensus around definitions and terminology, and
what exactly do we mean by initiation or use, or
whatever.

So I think in the spirit of future
directions, what would be your suggestion for us to
move collectively towards a consensus-building
around these terms and terminology? It's directed
to Dr. Levy, Georgetown University.

DR. LEVY: That's the million dollar
question. I think it's going to be very different.
I think we're going to need to look at different
measures of use and see what implications those
have.

I did some work with smokeless way back
that I'm extending now. And it seems like there
are clear patterns based on which measures you use.
But I think you've just got to look at the data and
see what predicts, see what gives you trends.

In the TUS, there were years where they asked, are you a regular user? And they asked, how many days in the past month did you use? I think that looking at those surveys, doing those kinds of surveys, can help give us some guidance as to what it means.

It becomes particularly difficult when you talk about initiation because how long does a person have to be using before they've initiated? It seems like a simple question, but I think it's really a tough question. And similarly about cessation, too.

DR. ROCHESTER: All right.

DR. MOYNIHAN: Michael Moynihan, Goodrich Tobacco. Yes, I would second that, that I think the definitions -- quantification of use is something that's been a real mystery in terms of looking at the data.

One of the things I'd like to compare it to is I recently somehow got on the Arbitron mailing list. I get mail from Arbitron constantly about radio surveys and TV diaries and all these

things and so on, which I don't know how I got on the list.

But I think in terms of different kinds of studies, these very large studies can't get to that level of detail. But we need some smaller studies that give us some sort of feeling for what's really happening in terms of patterns of use in a fairly rapid response kind of thing. And I don't know if that's among the studies that are already commissioned or not.

DR. CHOINIERE: I can address that. There are studies in planning to address some of those. But one of the key words that you said was "rapid." And if you're relying on a rapid study, the federal government is probably not the way to do it.

(Laughter.)

DR. CHOINIERE: So if researchers have their horn [indiscernible], we would encourage that type of research. But yes, we have both qualitative and some quantitative research planned to address, particularly, usage of some of these

new products. We don't even know how people talk

about use of some of these new products.

For instance, hookahs. How do people

talk about the use of water pipes? When you ask

them if you use it, what do they think? Are they

thinking about the one time they used it, or do

they think about -- and the same can be said for

all these various tobacco products. And there is a

more concerted effort among the agencies, FDA and

other agencies, to talk about how best to measure

some of these across all of our surveys.

DR. ROCHESTER: Go ahead.

DR. FLAHERTY: Brian Flaherty, University

of Washington. I just wanted to add, when talking

about measures of use and definitions of use,

Dr. Ware talked about psychometrics. I didn't

use the word, but I was also talking about

psychometrics. And often, measures work

differently in different subgroups, different

populations.

So establishing a definition in a

majority population may not apply to other

populations or other groups. So this definition of

measurements is probably going to be iterative over

looking at different populations, looking at

different hookahs, cigarettes, devices, whatever.

So the definition part might be complex.

DR. ROCHESTER: But necessary, however.

DR. FLAHERTY: But necessary, of course.

DR. ROCHESTER: But of course, to have a

collective conversation, I think we still have to

strive toward coming up with certain reasonable

definitions. Obviously, definitions are never

static. Right? We still have to iterate and

improve upon them with time.

But I think what I'm not hearing in terms

of future directions, though, is a way to

facilitate this. So we hear, well, maybe industry

needs to do something. But when we speak of

industry or regulators, we still need to have some

practical ways in which everyone is going to be

able to get into the sandbox and talk and play and

come up with some solutions. Right?

So any ideas on how to facilitate besides

saying, well, the FDA should do it? Go ahead.

(Laughter.)

DR. ROCHESTER: We'll come back to you.

DR. BOONE: Edward Boone from Virginia Commonwealth University. You've done a spectacular job of bringing all these modelers and putting them in the same room. And they're all competing against each other instead of working together.

You could easily facilitate this by initiating working groups. Even if you want to keep independence, have several working groups that pair up or coordinate with multiple researchers across academia, industry, the government. That way they're all coming up with -- they're sort of niched. Each of us has a niche.

But if we can work together, we can cross a lot of those boundaries instead of just saying, here's what I do. I do QALYs. And I don't do QALYs. But if I do QALYs, I could be useful to somebody who's doing population modeling because I can help them measure their endpoint.

By starting several, multiple working

groups that work -- I don't want to say

independently of each other or in competition of

each other, but are working to develop models,

you're probably going to get much more energy and

synergy than having everyone compete.

DR. ROCHESTER: I surely agree, and we've

used that model in other centers in other parts of

the government. Sure.

Go ahead.

DR. MEZA: There were some points I

wanted to add in terms of from the modeler

perspective, not necessarily FDA, but maybe all the

agencies, and just in general, whoever wants to

support the use of these models for policy.

So there are many things that could be

done. And I guess there are three points. One is,

support development of the diversity of these

modelers, or of these models. Right? So in some

sense, allow for, as was mentioned, the opportunity

for different approaches, different disciplines,

different perspectives to come together. Right?

And this workshop is a great example of that.

It doesn't end there, of course. There is funding -- I'm maybe speaking for myself -- and other modelers that probably would like to, of course -- and many of us are already funded in some ways to do this type of work. But of course, they need support for development of these models.

Also, support to get access to all the many data sources that are available. Many times it's very hard to get access to even publicly -- or, in theory, publicly available data that's already there, that already exists.

I'm not necessarily talking about new data that we would love to have and gather. It's access to existing data. And related to all of this, facilitate the synergies between modelers, discussions, interactions.

There are examples of a variety of networks, research networks that exist. CISNET is one of them that has been very successful in terms of producing results to help public policy, cancer public health policy. They visit the networks, MIDAS, infectious diseases.

So there are examples out there of successful collaborative networks of modelers that work together. Maybe some of them are more interactive, some of them are more independent, but at the end of the day, working towards a common goal and developing independent models trying to answer the same question.

This is the concept of comparative modeling, which I am biased because I'm part of a comparative modeling network. But I think it's something that has been very effective to give credibility to the modeling community in the public health.

DR. HAMMOND: I just want to echo that, that thinking about a formal network; that part of its purpose is to have sustained engagement between modelers, that that's part of the setup.

Over a substantial period of time -- both CISNET and MIDAS have been around for 10 years or more -- and I'm part of two other networks that are more recent; but that mechanism, where it's explicitly part of the funding and of membership in

the network to do this kind of ongoing regular talking and collaboration.

I would describe the relationship between modelers as somewhere in between collaboration and competition, or maybe both at the same time. But that can be very fruitful, actually, to have that mixture in the network.

DR. MEZA: So of course, then, what should modelers do, or is there any -- in terms of their actions, or things that we could do to make this even more successful.

So of course, models, in some sense, we talked about being parsimony. In general, models have to be practical, be designed to answer a question that's relevant. So in some sense, the modelers have to develop models that are responsive to the needs in this particular area.

We have to be more collaborative, at the same time competing among ourselves because it's a good thing; and, of course, the issue of transparency, right, developing models that are transparent, where the methods, the inputs, and the

outputs are clear, and through a variety of ways
that are already out there in terms of good
modeling practices.

Again, and reiterating some of the
comments, the issue of interdisciplinarity. And
many times, maybe, the agent-based modelers don't
want to talk with the system dynamics. Maybe
that's what we do because we belong in the same
type of goals.

But the issue of people working, quality
of life, or statisticians or more economics types
of areas, that's very fruitful. And that's where
these workshops are very helpful for all of us to
come together. But we have to be willing to make
that effort.

DR. ROCHESTER: Roger, and then Anne.

MR. PERKINS: Roger Perkins, Food and
Drug Administration but not CTP. And this is
solely my question.

First of all, I want to say this has been
a fantastic meeting. But when I look at it
holistically, I keep thinking to myself, all

behavior is based on biology. All biology is based

on chemistry.

So I want to address something a little

below the skin, as Dr. Hammond had mentioned it,

because I heard nothing about coupling models of

pharmacology or possible pharmacological

interventions with the population models and so

forth.

Because it seems to me there's at least

some probability that somebody's going to realize

either public funding, that there's a compelling

reason business of the huge cost burden of tobacco

products, or the private sector, that there's a

profit element there -- that somebody's going to

come up with something more innovative than chewing

gum, patches, and Chantix, which I guess has

undesirable side effects.

So I think NIADA says addiction is a

chronic relapsing brain disease. So the real

pernicious problem here is addiction, in some ways,

or that's a big part of it. And those are

neurobiological mechanisms. And it seems like if

you're going to build models looking out 5, 10, 20

years or whatever, that there's got to be some

probability that somebody's going to come along and

find some way to mitigate I think pretty potent

withdrawal symptoms and could change the trajectory

of all the models pretty radically.

DR. HAMMOND: That's an issue that looms

large in the food space, where there's a heavy

neurobiology aspect, and the characteristics of the

product matter a lot for it as well.

DR. ROCHESTER: Anne? Oh, okay. Anne

will be next.

DR. LIANG: My name is Qiwei Liang from

Altria. Dr. George Rochester, you start with a

very good presentation, basically set the tone from

your FDA perspective.

I'm thinking, can FDA build on your

presentation to develop some kind of protocol,

standard protocol, for describing model, whatever

the model the industry submit to FDA or whatever?

So follow a certain way to describe the model.

Remember, two things. Every model has

assumption. Many assumptions cannot be validated.

If they can be validated, they will become

parameters. So the assumptions cannot be

validated. So first you need to lay out what are

the assumptions for the model.

The second thing is about data.

Dr. Choiniere, the first thing you mention is data

collection need to be improved. That's very

important.

So in terms of data quality, there need

to be a description about the data. I have seen a

lot of the models with parameters based on poor

quality data. I think those estimates, they are

misleading. I would prefer not to have them.

So I think FDA should provide some

guidance. When you have very poor quality, should

there be a parameter estimate? For statistical

parameter, maybe you can give a confident interval.

Then people will have some kind of judgment. But

for some other model, because there is no confident

interval, then what to do?

DR. ROCHESTER: I think we've heard that

question to FDA several times about we should

strive, perhaps, towards providing guidance to

industry. And I think we've heard that, and we

will think about that and see how to address that

question. But there's always need for guidance in

every area of development, so we hear that.

Go ahead.

DR. CHOINIERE: You make a good point.

All models have assumptions. But one of the things

that George presented yesterday was the need to

document, and I believe Dr. Hammond as well, that

you need to document all of your assumptions and

the rationale for those assumptions, so that it's

clear to everyone looking at your model how you

came up with your result.

DR. ROCHESTER: I think the concept of

the guidance, though, in addition to that, is to

say for each person submitting their documentation

to FDA, what's the minimum threshold of quality and

level of detail and so on, acceptability, that

makes it reasonable in the context of what we are

looking for?

So I think you're asking us to communicate to the public in more clear terms, perhaps, some minimum standards for what we would expect from you in terms of documenting your model. So it's a reasonable request. I shall not commit, but I have heard you.

(Laughter.)

DR. CHOINIERE: But I think this is an area that was touched upon yesterday, and I think today as well. You have to test your model's assumptions, just do some sensitivity analysis around these assumptions to see how sensitive your model is, how important that assumption is to your ultimate result.

If the results are highly contingent on this one parameter that has a large amount of uncertainty around it, I think that then it makes it harder for you to justify that model, I would assume, for the purpose that you are proposing.

DR. ROCHESTER: But in general, we agree that we have some overriding principles that, across almost any type of model that you're going

to develop and present, that these principles would
be best practices that we can actually agree on and
I think we could implement.

We're all so quiet. It's late in the
evening, and we want to go home, and we want to get
a flight. But I'd like to get a little more life
going here. So I would like to say I haven't heard
anything at all about lack of data in terms of
understanding youth behaviors, youth initiation,
et cetera.

So in the situation where not a lot of
studies would be done on youth or haven't been
done, and you have to look at historical data or
you don't have enough data, what are the main
approaches that one thinks of in terms of trying to
understand youth within the culture of this work
that needs to go forward?

Dr. Backinger, you're smiling. You're
like, why did he say that?

No. I just think in youth, where we
don't have a lot of data and we look at the future,
we will probably have even less data in the future,

but we still have some questions we need to answer.
So perhaps we could -- I haven't heard anything at
all about that.

MS. LEE: I'm not exactly sure -- Monica
Lee from JTI, by the way -- but I was thinking very
similar along the line when you were bringing up
that point, because when the session started
yesterday, Dr. Zeller started with, as an example,
two products and how it can be complicated. And
then he is hoping for a lot of discussion.

I also know, as many of us here, that
tobacco modeling has been going on well before we
started talking about product, especially from
tobacco control/public health perspective, since
NCI had that monograph. So that has a lot longer
history and understanding.

So me coming from industry, everyone is
thinking about what data do we need to generate,
what kind of model we have to generate to submit
the MRTP and application. And Dr. Rochester's
really nice outlines about what you may be thinking
in terms of a qualification, that was very helpful.

At the same time, and I was listening to a lot of current, existing example models, I can't help but thinking how many models can actually meet the criteria. Because a lot of -- I'm not pointing out any specific model -- because a lot of those tobacco control models are geared toward the advocacy messages. Therefore, point estimate was okay for making some points.

But when you're actually applying to a regulatory framework, as you said, validations and cross-check, I don't think that many models can meet that quality.

I think Dr. Levy has, I think, probably among many other models, the most transparent, relatively speaking. He always has supplemental data, and I think many people appreciate that. But we don't see that kind of transparency in other models.

So one point I think I bring up to everyone involved here is that not just for the industry perspective, trying to meet the quality. Anybody bringing the tobacco modeling should keep

that in mind and they should meet that.

My second point, I think more relevant to Dr. Rochester's point, is me coming in here, what I was -- high expectation is related to the product. And we cannot really generate 5, 10-year survey data for one product.

So having that, we are looking for short-term clinical, perhaps, perception study, opinion studies, what they may do, may not do. And that kind of example data set has not been discussed. And I was hoping in the future, if you have more discussion, I would very much be looking forward to it.

DR. MEZA: Can I make a comment about the first point there about model quality? That's one of the things or one of the issues where having a collaborative network is really helpful. And I can tell you about the experience in CISNET, where CISNET has really pushed modelers to be extremely transparent about their assumptions.

We have model profiles that are publicly available that are in standard format so you can

look at all the CISNET models and read them from
the same perspective, comparing the different
inputs, the different outputs, and with the same
notation and terminology and all that.

So this is something where these
collaborative networks are very helpful to make
sure that there is a uniform minimum level of
quality in model development.

Maybe just quickly going back regarding
the comment about data on youth, as modelers, we
face this issue all the time. And of course, we
have to get creative, and maybe that's an example
of the analysis that Ted Holford showed yesterday,
right, where you're using cross-sectional data that
asked about initiation to get a sense of what was
happening when these adults that you surveyed were
youth, were young.

So it goes back to the issue like many
times, not for all health aspects -- I don't know
the things that are relevant -- we actually ask
about, when did you start this behavior. And
that's something that's very helpful. And for

tobacco, for smoking, it has been really helpful to characterize initiation during youth, to have information that you asked later in life.

The other issue is what Ross mentioned earlier, which was the use of these new forms of data like social networks, Facebook, Twitter, that somehow are probably going to help us try to get some information about the patterns of behavior of youth, where it's not the usual way that we'd get information. Right? But young people are putting a lot of information out there, and I guess it's up to us to go and take it.

DR. ROCHESTER: Did you want to add anything?

DR. HAMMOND: I'll just say that I very much appreciate what you were saying. And it's not only that collaborative networks can impose some structure and clarity on models and assumptions, but that there's economies of scale in a network in obtaining access to data and in making information about models public that are onerous for any single modeler, but for a network can be more easily done.

DR. ROCHESTER: All right. Anne? And she's been burning for a while here.

MS. HARTMAN: I just think the idea of networks -- oh, Anne Hartman, NCI -- a network, or working groups, however you want to say it, is really the most relevant. But you've been emphasizing the modeler part and the consistent data to understand the models. And that's one aspect, and that's very good for validity.

But I want to go back to what was said yesterday in the beginning, and probably today several times, and Ross also. There's a whole big thing in deciding on that and the inputs.

So you need to have, as part of that network, the people who collect data or have data or know about the assumptions of the data they're collecting and how much you can utilize that. So you've got to remember the front end.

DR. MEZA: I completely -- I can give you an example of CISNET lung, or actually, it's group of modelers that is complimented with people from NCI, experts. In this case, it's lung cancer

experts, maybe clinicians, maybe people that have other types of expertise.

The example I wanted to mention is a recent collaboration that we had with the U.S. Preventive Services Task Force to look at lung cancer screening recommendations. And to do that, we had to get access to the two large trials, which is NLST and PLCO.

So going back to the economies of scale, we got access because it was a collaborative network and NLST and PLCO were willing to work with us. And then when we got the data -- it's not that we just got the data, and then that's okay, goodbye, and then we'll come back. No.

Actually, two or three members of NLST and PLCO became active members of the project and the groups that we were doing. And actually, that's what made it possible. Right? Because they really helped us work through the nuances of the characteristics of the data, exactly trying to understand why is this like this, why is this like that. And you had the person right there all the

time to ask them.

Now that we are done and the papers that are presenting our results of the effectiveness of CT screening for lung cancer, you'll see that the list of co-authors include the people of NLST and the people of NCI and many others. It's not only the modelers. It's the whole -- and that's actually so true.

DR. HAMMOND: Let me just add one thing to that, which is whether it's in a network or not, I think there's an important point implicit in what you said that has been hinted at a couple times throughout this two days, but I want to highlight it again here, which is that the ideal relationship between data and modeling is not linear but is cyclical.

Ideally, models not only are consumers of data but they help inform what data should be collected and the priorities for data collection, and even the form and the methodology that might be most useful for that data to be collected in the first place.

So in some of these networks and working
groups, and sometimes in just freestanding
communities of people who do modeling, there are
these ongoing relationships over very long periods
of time where there's back and forth between models
and data collection over and over, and that's how
you build up very complex and rich efforts.

MR. FINLEY: Patrick Finley from Sandia
Labs. Just following up on something that Ross
just mentioned, I think one of the interesting
threads from yesterday was this very topic, that
models often will tell you the kind of data that
you need to make advances.

I believe it was Gary Giovino that
mentioned briefly the idea of rapid turnaround type
of data. That's something I'd like to just throw
out for us to think about in the longer term,
because I can consider a number of times when we
find those surprises in the model run. And it
would be really great to see almost data to order,
if you will, to actually be able to evaluate, is
this an artifact or are we actually looking at

something?

I know a number of colleagues modeling in other fields who reluctantly rely on things -- for example, Mechanical Turk surveys, situations where you consciously sacrifice accuracy and accept a certain amount of bias just for the quick turnaround to at least be able to get off your high center and make some progress during the time that you're actually waiting to develop and validate a survey that would give you the true data.

So I think that, as a group, it would be interesting to think about how we can have, if you will, a sliding scale of data quality, almost, something that you can get rapid turnaround to be able to continue your work. Hopefully during that time you can get better data that will enable you to get truly validatable types of results.

DR. SARKAR: Mohamadi Sarkar from Altria Client Services. I want to go back to something that -- you posed a provocative question about how are we going to get youth to -- before I address that, I want to thank CTP for organizing this

workshop. It was phenomenal, and I hope that you continue to engage in this three-way dialogue.

That's a question that keeps us awake at nights as well. So as we think about developing MRTPs and new product applications, how are we going to address the question about initiation, specific for the product under consideration?

For example, you might have data on the category. Let's think in e-cigarettes, if there is information in the category for initiation out from surveys or even these rapid response experiments; but the ethical considerations of asking youth about perceptions and usage of a new product before it's introduced in the market, and how do we gather that data with enough confidence that the modelers can use it, and perhaps in these kinds of situations, a what-if scenario might give some directional evidence.

Even if that's done, at least the assumptions should be reasonable. You don't want to create a what-if scenario that's completely unrealistic. But some realistic assumptions and

creating what-if scenarios, I'd like your thoughts
on that.

DR. CHOINIERE: Can you clarify what you
mean by the what-if scenario? Are you
saying -- can you just explain what you're talking
about?

DR. SARKAR: So let's consider a brand
new product for which you have no information on
initiation. One potential option to address
Dr. Rochester's question on how do you gather
information on initiation to conduct a study in
youth and do some susceptibility or perception type
of assessment, without taking it any further, you'd
get some responses, but that may not necessarily be
directly reflecting what they'll actually do in
terms of behaviors. You might get some opinions.

But in the absence of that, another
possible scenario could be that you could create a
model and then put in some various hypothetical
scenarios of 1 percent initiation, 5 percent
initiation, and so on and so forth. And perhaps
that might be one way of assessing the overall

impact on the population.

DR. ROCHESTER: So you're just saying
using some possible values from perhaps historical
experience with other products, perhaps.

DR. SARKAR: Right.

DR. ROCHESTER: But you are saying not to
have wildly crazy assumptions that we know
60 percent of youth won't initiate or something
like that.

DR. CHOINIERE: I certainly think there's
value into doing these scenarios, worst-case
scenarios, seeing how bad can it get. There's a
lot of talk about tipping points over the past
couple days.

One approach I could see is that as
opposed to trying to predict what the level of
adoption might be or the level of initiation among
youth may be, is find out what the range of values
are where, given the likely health effects of the
product, what's the range of values for all these
different types of behaviors over which it would be
appropriate for public health.

What are the red zones that we need to be worried about? Are there certain levels of initiation among youth that we really need to start being concerned about? So that way, if we know -- I guess that would be like a surface plot of some sort of all these various possible transitions.

If we know that set of transitions, then when we are doing surveillance and we see things going outside of that range, then we know perhaps we have a problem that we need to address. I'm just suggesting that that's a possible way.

DR. ROCHESTER: Right. So there's a clear role for projecting, predicting, and --

DR. SARKAR: Validating.

DR. ROCHESTER: -- validating, and surveillance to confirm or to monitor. Right. Surveillance, of course, is another topic of interest.

Any other burning points?

(No response.)

DR. ROCHESTER: All right. This meeting

was called I guess with a clear objective to focus

on the scientific considerations in terms of

developing models and minimize discussions on

policy, which are just never avoidable.

I think by having us talk more about the

science in a general framework without the context

of making any decisions, I think that allowed us to

have a richer dialogue today.

However, it's always burning in

somebody's mind because they want to hear exactly

what decision you're going to make if I give you

this information. I know the temptation always is

to ask this kind of thing.

Then it's my responsibility, when I am

chairing the session or moderating, to certainly

not fall for the temptation to answer such

questions. I do not claim that I'm avoiding your

questions; I will claim that I'm postponing the

answer to another time.

It was meant to bring together different

stakeholders, and I think here we have regulators,

regulated industry scientists, academic

researchers, people from the tobacco control

community, and other stakeholders, which has

enriched our discussion. But clearly, there are

other groups that perhaps were not able to make the

meeting at this time. For instance, we don't have

members from, let's say, CDC, which is our largest

disease control and prevention agency, and so on.

So we should think of the future to

perhaps allow a little bit more time in our

planning to be able to accommodate other people

that perhaps were not at this meeting.

Clearly there is rich dialogue, in my

mind a good, respectful tone that allowed everyone

to share opinions and ideas and be able to think.

And I feel like we have a collective strategic

energy that has developed.

Certainly the challenge from here, then,

is to build up on this going forward. There's

obviously some role for FDA or CTP in terms of

leadership, facilitating some of these things to

happen. And there's a role for the non-CTP type

community, who need to contribute in different

ways, not just financially, but also in terms of
the talent and experience, et cetera.

Sharing of tools becomes important
because until we're able to fully understand and
exploit the opportunities and the challenges that
surround the use of any of these models, or the use
of many of them, to enrich our decision-making
ingredients -- I think that until we've got that
richness that really helps inform regulators of the
breadth of experience and the breadth of models
that exist or needs to be developed, I think that
will make for better public health in general.

Clearly there are some gaps that need to
be filled. And so in terms of the research going
forward, that's being sponsored either by CTP, the
NIH, industry, and others. And as you're thinking
of developing your portfolio for marketing any kind
of product, I think there's a rich set of questions
that now we can go forward thinking about that
certainly needs some adaptation to your own
specific needs in the context of your own
application.

Before I close the meeting, I always leave this thing for the last, and I usually forget. I always say goodbye, and then I go, oh, I never thanked everyone that worked so hard. So first upon the list you need to thank George Rochester for a great job.

(Laughter.)

DR. ROCHESTER: I was just teasing. Just kidding, just kidding. But our great scientific program committee; I want to acknowledge the great inputs we received from Ben Apelberg.

(Applause.)

DR. ROCHESTER: Dr. Carolyn Dresler.

(Applause.)

Adjournment

DR. ROCHESTER: Eric Backlund; Danny Lee; Nicholas Farris; Brian Rostron; Charles Wu. And our advisors David Ashley, Kathy Backinger, Conrad Choiniere. Our advisors and consultant staff, so Karen Sommers and Caryn Cohen and others. And our technical staff that have just made this so far, I think, a lovely meeting.

At this time I wish you safe journeys, a wonderful flight, and if you're going to Arkansas, please wait till tomorrow. Do not go to Arkansas tonight. It's all iced over.

So at this time, I would like to close the meeting and thank you all for your wonderful participation.

(Whereupon, at 4:26 p.m., the meeting was adjourned.)